DOCUMENT RESUME

ED 298 449 CS 009 300

AUTHOR Powell, Janet L.
TITLE An Examination of Comprehension Processes Used by Readers as They Engage in Different Forms of Assessment.
PUB DATE Dec 88
NOTE 196p.; Ed.D. Dissertation, Indiana University.
PUB TYPE Reports - Research/Technical (143) -- Dissertations/Theses - Doctoral Dissertations (041)

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS *Cloze Procedure; Construct Validity; Elementary Education; Grade 6; Multiple Choice Tests; *Reading Comprehension; Reading Research; *Reading Strategies; *Reading Tests

ABSTRACT

A study described and compared the reading and thinking processes used by subjects while engaged in a multiple choice test, a cloze test, a written retelling, and a nonassessed passage to see whether different strategies were used for different forms of comprehension assessment. Subjects, nine teacher-identified proficient sixth grade readers, completed three multiple choice tests, three cloze tests, three written retellings, and three nonassessed passages and were interviewed for each assessment task. The data consisted of the subjects' hand and eye movements observed while reading and their verbal protocols given as they used concurrent and retrospective introspection. Results indicated that there were at least some superficial differences in the way subjects engaged in each task. A categorization of 21 reading processes which emerged from the verbal reports was devised. Results also indicated that the cloze test was perceived as the most different from the other tasks and that it may lack construct validity. (Thirty-five tables of data and six figures are included; 112 references and an appendix of data are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED298449

AN EXAMINATION OF COMPREHENSION PROCESSES USED BY READERS AS THEY ENGAGE IN DIFFERENT FORMS OF ASSESSMENT

Janet L. Powell

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Janet Powell

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Submitted to the faculty of the School of Education
in partial fulfillment of the requirements
of the degree
Doctor of Education

Indiana University

December, 1988

CS009300

BEST COPY AVAILABLE

Accepted by the School of Education, Indiana University, in partial fulfillment of the requirements of the degree of Doctor of Education.

Doctoral Committee:

Roger Farr, Chairperson &
Director of Thesis

Leo Fay

Richard Pugh

Harmon Baldwin

October 17, 1988

Copyright by

Janet L. Powell

1988

ACKNOWLEDGMENTS

Above all, I would like to express my sincerest gratitude to Roger Farr, whose inspiration, encouragement, and guidance were instrumental from the conceptualization to the completion of this research.

I also want to express my appreciation to Leo Fay, Rick Pugh, and Harmon Baldwin for their advice at crucial points of the study. I also deeply appreciate the help and support of all of the graduate students who assisted me in various ways with the dissertation, especially Brian Smitten, Caroline Beverstock, Jamie Myers, and Bruce Robbins. In addition, I wish to thank the two elementary teachers who allowed me to take students from their classrooms and the students themselves who participated in my study.

Finally, but not least, I wish to thank my parents, James C. and Agatha D. Powell, and my children, Doug, Jessica, and Justin for their patience, inspiration, and understanding.

TABLE OF CONTENTS

Page

LIST OF TABLES..... vi

LIST OF FIGURES.....viii

CHAPTER I: INTRODUCTION..... 1

Background and Rationale..... 1
The Active Role of the Reader..... 4
Reading Comprehension Tests..... 8
Overview of Study..... 13
Summary..... 46

CHAPTER II: A REVIEW OF THE LITERATURE..... 15

Cognition and Reading Comprehension..... 16
Verbal Reporting of Reading Processes... 22
Reading Comprehension Test Validity..... 30
Summary..... 46

CHAPTER III: METHODOLOGY..... 47

Pilot Study..... 47
Subjects..... 49
Materials..... 52
Data Collection Procedures..... 55
Data Analysis Procedures..... 76
The Qualitative Analysis..... 77
The Quantitative Analysis..... 82
Limitations..... 86
Summary..... 90

CHAPTER IV: RESULTS..... 91

The Qualitative Results..... 92
The Quantitative Results.....114
Summary.....142

CHAPTER V: CONCLUSIONS.....144

The Qualitative Data.....145
The Quantitative Data.....154
Summary of Conclusions.....165
Implications.....169

REFERENCES.....173

LIST OF TABLES

Table		Page
1	Nature of Passages.....	54
2	Data Collection Procedures.....	61
3	Cognitive Processes by Subjects.	104
4	Cognitive Processes by Tasks....	107
	Cognitive Processes by Subject	
5	Subject 1.....	117
6	Subject 2.....	118
7	Subject 3.....	119
8	Subject 4.....	120
9	Subject 5.....	121
10	Subject 6.....	122
11	Subject 7.....	123
12	Subject 8.....	124
13	Subject 9.....	125
14	Totals.....	126
	Cognitive Processes by Process	
15	Recall.....	128
16	Key Word.....	128
17	Eliminate.....	129
18	Read Ahead.....	129
19	Spell.....	130
20	Guess.....	130
21	Judge Importance.....	131

22	Main Idea.....	131
23	Organization.....	132
24	Skip.....	132
25	Context Clues.....	133
26	Reread.....	133
27	Paraphrase.....	134
28	Prior Knowledge.....	134
29	Opinion.....	135
30	Predict/Form Questions.....	135
31	Confirm/Disconfirm.....	136
32	Vary Rate.....	136
33	Fail to Understand.....	137
34	Speculate.....	137
35	Visualize.....	138

LIST OF FIGURES

Figure		Page
1	Outline of Procedures.....	63
2	Observation-Only Interviews....	94
3	Test-Related Processes.....	96
4	General Reading Processes.....	100
5	Responses to Exit Interviews..	109
6	Results of Post-Hoc Analysis..	141

CHAPTER I

INTRODUCTION

Background and Rationale

The assessment of reading is an important and often controversial topic. The public and private sectors are demanding more accountability through standardized testing from the public schools. Often this accountability is based on the administration of standardized tests. It was estimated in 1954 that 75 million standardized tests were administered to students (Kirkland, 1971). In 1982, Anderson reported the following:

> In 1975 Houts estimated that in the United States each student receives from six to twelve full batteries of achievement tests during the years from kindergarten through high school. This estimate did not even take into account specialized achievement testing, locally developed diagnostic tests, testing done through the National Assessment of Educational Progress, or competency tests now in effect in many states. Given a 1978 population of about 48 million 5 to 17 years old, the number of tests administered to elementary and secondary school students each year must be in the hundreds of millions. (p. 232)

By 1981, the estimates of the numbers of tests administered had risen to 400 to 500 million (Strenio, 1981). Despite this significant increase in test usage, many reading experts are questioning the very essence of standardized reading tests by claiming that they are an invalid measure of reading comprehension. The long list of criticisms include test bias, over-dependence on statistical interpretations, lack of breadth and depth of content, ambiguity, penalties for deep thinkers, lack of diagnostic value and information gain, test-wiseness, and forcing readers into a search for one correct meaning (Hoffman 1962; Houts, 1977; Strenio, 1981; Fransson, 1984).

As reading educators concerned with the assumed discrepancy between what is known about reading comprehension and its measurement attempt to understand this hypothetical discrepancy, political and public pressure continues to require the assessment of reading. Reading educators do not have the luxury of halting the assessment of reading until more research about reading comprehension is completed, and more valid ways to assess it are developed. The standardized assessment of individual readers should always be considered as only one part of a reader's total assessment, However, standardized tests continue to be a very significant

factor since many state governments and local school districts require by law or policy that standardized tests be used for school accountability.

Validity is the most important aspect of any test and it is the validity of reading comprehension tests that is being questioned. Construct validity is of particular importance. Construct validity is concerned with whether a test measures the behaviors it is supposed to measure--in this case reading comprehension. All assessments of reading are indirect in that we cannot actually see the processes involved. Assessments of reading can only measure the products of reading. These products are the answers given on tests, which are turned into scores which alone tell us nothing about the process. By examining the processes involved in taking various reading tests and comparing these processes across assessment tasks, we may be able to gain some clarity as to how the reading processes differ. This would add some knowledge to the questions concerning the construct validity of reading comprehension tests. For example, the reading processes may be significantly different as subjects take a multiple choice test as they would if the subject were reading a nonassessed passage. If this were true, then the construct validity of multiple choice tests could be challenged. This

study examined the comprehension processes readers used while engaged in a multiple choice test, a cloze test, and a written retelling. For purposes of comparison, a nonassessed reading task was also included. Specifically, the study focused on the following question:

> Do readers use different reading/thinking processes dependent on the form of assessment used?

The Active Role of the Reader

While educators may argue about the best way to teach reading, all agree that comprehension is the ultimate goal. Grapho-phonic, syntactic, and semantic information are all important text clues that lead to comprehension (Goodman, Y. & Burke, 1971; Goodman, K., 1967). However, while text clues such as these are an important part of reading comprehension, researchers have for a long time known that there is a great deal more to reading comprehension than a skill or sequential model of reading suggests (Thorndike, 1917).

Devine (1986) defines reading comprehension as:

> ...the process of using syntactic, semantic, and rhetorical information found in printed texts to reconstruct in the readers' mind, using the knowledge of the world he or she possesses, plus appropriate cognitive skills and reasoning ability, a hypothesis or personal explanation which may account for the intended message that existed in the writer's mind as the printed text was prepared. (p. 67)

The active role of the reader in the reading process has recently gained much attention in reading research. We have learned that proficient readers are strategic as they use print clues to construct the meaning of a text (Afflerbach and Johnston 1984; Baker and Brown 1984; Chaffee 1985; Cohen, 1986). In addition to the text, proficient readers use background knowledge of the subject matter, prediction, comprehension monitoring and repair strategies, and imagery, as well as their knowledge of the language to reconstruct meaning (Collins, Brown, and Larkin, 1980; Pearson and Spiro, 1980). The ability to be flexible and vary reading strategies to fit specific reading tasks is what makes a reader proficient.

The role of the active reader has become widely accepted by reading experts. In a position paper of the Michigan Department of Education and the Michigan Reading Association, Wixson and Peters (1984) asserted that "reading is the process of constructing meaning

through the dynamic interaction among the reader, the text, and the context of the reading situation."

Attempts to identify and define reading comprehension strategies, also known as metacomprehension is relatively recent (Anderson & Armbruster, 1982). Van Dijk and Kintsch (1983) define a strategy as "the idea of an agent about the best way to act in order to reach a goal" (64-5). Spiro (1980) views comprehension strategies as "any deliberate, planful control of activities that gives birth to comprehension" (p. 456). Rowe and Rayford (1985) define reading strategies as "purposeful actions taken voluntarily to achieve particular outcomes". The nature of this study is to examine how certain reading tasks affect reading behaviors. Because Rowe and Rayford's definition includes the recognition that readers have particular outcomes in mind as they read, this definition of reading comprehension strategies will be used. However, this study will also include mental processes which may not be deliberate or planful, such as tying prior experiences in with the text and imagery. Other research studies examining reading processes have included such things as imagery as a reading <u>strategy</u>. It is felt that the occurrence of such things as imagery may not be deliberate and planful, yet may be

significant for proficient reading to occur. Therefore, this study will use the term "process" in place of "strategy". The term process as used in this study will include reading strategies as defined above. Therefore, all thoughts expressed by the subjects will be considered a part of the reader's comprehension processes.

It seems reasonable to suggest that reading processes may change somewhat depending on the task and the purpose for reading. It is possible, for example, that readers skim a passage more often in taking a multiple choice test than they would if they were reading the text for another purpose. It could then be inferred that the reading processes differed because of the task. If these differences are significant, the answers to some of the construct validity questions regarding reading tests may become clearer. By understanding the differences in reading processes the assessment task may cause, we may better understand the construct validity of various types of reading comprehension test tasks.

Reading Comprehension Tests

Multiple choice tests are the most popular forms of reading comprehension assessment. One of the reasons why multiple choice tests are so widely used is because of the ease of scoring them. Cloze tests and written retellings are also popular measures of reading comprehension. The reason they are not as popular as multiple choice tests is probably related to the fact that they are not as easily scored. The following section will provide a brief description of each of these tests, what they claim to measure, and some of the criticisms associated with each.

Multiple-Choice Tests

The majority of standardized reading tests are multiple choice. Farr (1986) states that "the manuals of most standardized [reading] tests make very explicit the fact that the test will not provide information about a pupil's reading processes but only information about the product of reading". However, he continues by saying that "...one could argue that the product-or score- isn't valid if a pupil doesn't use the actual processes of reading in determining the answers." Despite the fact that most test manuals state clearly

that multiple choice tests do not measure the process, only the product of reading, and that the results must be considered within the limitations of the test, the criticisms surrounding them abound.

In the introduction to an edition of The Reading Teacher that focused on the state of reading assessment, Valencia and Pearson (1987) stated:

> The tests used to measure reading achievement do not reflect recent advances in our understanding of the reading process. If we are to foster effective instruction, the discrepancy between what we know and what we measure must be resolved (728).

Cloze Tests

The strict cloze technique has been used in reading instruction and assessment for several years. The technique forces students to use the context of a passage to suggest replacements for deleted words. In the first description of the cloze technique, Taylor (1953) describes the procedure as:

> ...a method of intercepting a message from a "transmitter" (writer or speaker), mutilating its language patterns by deleting parts, and so administering it to "receivers" (readers and listeners) that their attempts to make the patterns whole again potentially yield a considerable number of cloze units (416).

The Degrees of Reading Power (DRP) is a standardized cloze test. The DRP Handbook (1986)

claims that the tests are "holistic measures of how well the messages within the text are understood." The handbook offers the following explanation of what cloze tests measure:

> DRP tests do not assess the products of reading-the ability to go beyond the message in the text to make, for example, evaluative judgments regarding the message. They do measure what might be called minimally inferential comprehension-a necessary, though not sufficient prerequisite to other higher-order cognitive abilities (2).

The handbook claims that "it is not possible to answer test items correctly by relying only on the information in the sentence containing the blank. A paragraph, or at least several sentences, must be understood to respond successfully."

The DRP tests are modified cloze tests. The correct response and three distractors are provided for each blank. A strict cloze test was used in this study. Nevertheless, it seems reasonable to assume that both strict and modified cloze tests measure the same thing. Readers must rely on clues within the text to determine which words will fit into the blanks. Goodman (1967) describes three major text clues as (1) syntactic or grammatical word order clues; (2) semantic or meaning clues; and (3) graphophonic or sound-symbol clues.

Cloze tests have been criticized on several grounds. Ashby-Davis (1985) argues that cloze reading

is not like ordinary reading. She states that reading speed, eye movements, and overall reading strategies are changed during cloze testing. She also believes that strict cloze tests are not a valid measure of reading comprehension because they measure more than reading comprehension, that being the ability to write.

In this study, a strict cloze test will be used since modified cloze tests vary greatly and many are very similar to multiple choice tests in that the students are to select one correct answer from a set of four or five choices.

Written Retellings

Retellings of stories, having students write or tell you everything they can remember about a story, have been used for many years to gather data (Kalmbach, 1986; Bartlett, 1932). Retellings have been used extensively as an assessment tool in reading research, but they have not been used as commonly in the classroom. Smith and Jackson (1985) state that retellings "give us a sense of how as well as how much information is represented in the student's thinking right after reading" (622). Kalmbach (1986) claims:

> Specifically, retellings can reveal: (1) The point or points students see in the stories they read; (2) The problems students have organizing the different elements of a story into a coherent

> whole. Armed with such knowledge, teachers can evaluate the effectiveness of the stories they are having students read as well as gauge the impact of those stories on students. In addition, analyzing retellings can reveal the strengths of weaker readers and the difficulties stronger readers can sometimes have manipulating texts (327).

Written retellings have been criticized for not only measuring how well students read, but also how well they write. However, most of the criticisms surrounding retellings relate to the manner in which they are scored. Most methods assign points to reflect the relative importance of elements such as character, plot, and setting (Clark, 1982; Goodman and Burke, 1971; Smith, 1980). Critics of the scoring methods state that "systems for evaluating retellings should incorporate the balance among aspects of the text and the effect of the whole" (Irwin & Mitchell, 1983) and that retellings should be analyzed by their independent structure rather than comparing them to the original story (Kalmbach, 1986). However, many of the criticisms surrounding the scoring of retellings are lessened when a carefully defined method of scoring that takes many of these criticisms into account is implemented.

Overview of Study

This study is descriptive of the comprehension processes readers employ given three forms of reading assessment tasks and a nonassessed reading activity. It also served as an examination of the processes used by those subjects who scored at a high level of comprehension according to the traditional scoring of each type of test. Over the course of twelve interviews given on different days, nine proficient sixth grade readers were given a total of twelve tasks. The tasks consisted of three multiple choice tests, three cloze procedures, and three written retelling procedures. They were also asked to read three passages without any formal assessment to serve as a comparison to the assessed reading tasks. The data consists of their eye and hand movements while being observed and the subjects' verbal protocols given by the subjects as they used concurrent and retrospective introspection. During the first four sessions, the subjects were only observed as they completed each task. Thereafter, subjects used introspection and retrospection to describe what they were doing as they completed each task. The use of introspection as well as the tasks and passages were rotated to control for order effects.

Summary

This study yields reasonably valid and reliable data on how reading processes differ as readers engage in multiple choice tests, cloze tests, written retellings, and a nonassessed reading activity. As a result, the findings provide insights into the construct validity of the three reading tests as well as implications for further research.

CHAPTER II

A REVIEW OF THE LITERATURE

This chapter is divided into three sections. This study is an investigation of the thinking processes associated with various means of assessing reading comprehension. Therefore, section one provides a theoretical framework of cognition and relates it to reading comprehension and the metacognitive awareness of the reading process. The methodological concerns of collecting data on reading processes are vital to the validity of this study. Consequently, section two reviews the critical issues concerning the validity of introspection, the data collection procedure which was used in this study. In addition, this research is an examination of how reading tests alter reading processes, which raises important questions concerning the validity of the reading assessment measures. Hence, section three concerns the debate over the validity of reading comprehension assessment and critically reviews those studies which have examined the reading processes under testing conditions.

Cognition and Reading Comprehension

Thinking is a cognitive process, a mental act by which knowledge is acquired (Presseisen, 1984) and manipulated (Cohen, 1971). These processes require an active involvement on the part of the thinker. Complex relationships are developed during this process which can be interconnected as an organized structure by the thinker in a variety of ways. The following definitions of thinking illustrate some of the complexities of the thinking process:

> Thinking is defined as the mental derivation of mental elements (thoughts) from perceptions and the mental manipulation/combination of these thoughts (Cohen 1971).
>
> Thinking is the mental manipulation of sensory input to formulate thoughts, reason about, or judge (Beyer 1984).
>
> Thinking can be defined as: the extension of evidence in accord with that evidence so as to fill up gaps in the evidence: and this is done by moving through a succession of interconnected steps which may be stated at the time, or left till later to be stated (Bartlett 1958).

The literature on thinking provides multiple lists of cognitive processes that are considered thinking skills. Bloom (1956) and others (Guilford 1967; Flavell

1976) define thinking in terms of hierarchies of questioning associated with critical thinking theories.

Each of Bloom's cognitive categories includes a list of thinking skills. They are expressed as objectives or goals of specific learning tasks. For example:

Knowledge:	define, recognize, recall identify, label, understand, examine, show, collect.
Comprehension:	translate, interpret, explain, describe, summarize, extrapolate.
Application:	apply, solve, experiment, show, predict.
Analysis:	connect, relate, differentiate, classify, arrange, check, group, distinguish, organize, categorize, detect, compare, infer.
Synthesis:	produce, propose, design, plan, combine, formulate, compose, hypothesize, construct.
Evaluate:	appraise, judge, criticize, decide.

There are many arguments against such categorizations of thinking processes. Beyer (1984) argues that these and other such categories are often ambiguous and difficult to distinguish. Beyer also criticizes such works because the tasks generally move from simpler to complex operations, from more observable and concrete to abstract dimensions, and from an emphasis on working with known materials toward creating

or inventing new, previously unknown approaches or materials. These hierarchies imply that one category is superior to another, and suggest that there is some sort of orderly progression from one level to another. Thus, it can be implied that these hierarchies are based on the assumption that learning is linear or sequential. Nickerson (1981) suggests that no one taxonomy exits and that educators would be wise to select abilities that represent what they want students to be able to do and incorporate these particular skills into their curricula. In addition, most models of thinking are based entirely on the rational or logical mind, and overlook the affective or emotional aspects of human thoughts.

While there may not be any one taxonomy that scholars can thoroughly agree upon, many have the same basic premise, that there are differences between lower, essential skills and complex, multiple process-strategies that involve numerous skills. Although the terms used in Bloom's and other such hierarchies are often ambiguous, they are useful in discussing thinking. Therefore, if Bloom's terminology is used, but put into a model that doesn't suggest separateness and hierarchies, and the affective domain is included, the problems are at least partially resolved. This model

can then serve as a useful framework for a definition of thinking.

A Model of the Thinking Process

Knowledge

Application Analysis

Comprehension Synthesis

Evaluation Affect

The thinking process is based on the knowledge of concepts, and includes the evaluation, application, analysis, comprehension, synthesis, and feelings (affect) about these ideas. It seems reasonable to suggest that the significance of any one of these factors upon the process is dependent on the topic of the thinking and the goal or purpose of the thinker.

In 1917, Thorndike defined reading as reasoning. Reading is a cognitive process. Reading is thought guided by printed symbols (Farr & Roser, 1979). Reading involves actively thinking about a text as it is being read. Readers construct their own meaning and attitudes of print as they apply, evaluate, analyze, and synthesize the meaning based on their prior knowledge

and experience. Readers consolidate information from a text with what they already know and feel about the subject. This interaction between the reader and the text allows the reader to alternate between prior knowledge and the text to gain meaning (Singer, 1983; Adams & Collins, 1977; Rumelhart, 1976, Winograd, 1972).

The manner in which the mind observes its own process can be traced back to Aristotle and Plato (Boring, 1953). A reader's awareness of thought processes involved in reading has recently come to be known as metacognition. Flavell (1976) defines metacognition as the awareness of, and an ability to capitalize on, one's own knowledge and thought process as these are applied to some specific task. The general knowledge of the reader guides him or her in monitoring comprehension processes through the selection and implementation of specific strategies to achieve some predetermined goal or purpose for reading. The chief idea involved in metacognition is that learners must actively monitor their use of thinking processes and regulate them according to the purpose for reading.

Presseisen (1984) describes two main dimensions to metacognitive thinking. The first is task-oriented. This dimension has to do with monitoring the actual

and experience. Readers consolidate information from a text with what they already know and feel about the subject. This interaction between the reader and the text allows the reader to alternate between prior knowledge and the text to gain meaning (Singer, 1983; Adams & Collins, 1977; Rumelhart, 1976, Winograd, 1972).

The manner in which the mind observes its own process can be traced back to Aristotle and Plato (Boring, 1953). A reader's awareness of thought processes involved in reading has recently come to be known as metacognition. Flavell (1976) defines metacognition as the awareness of, and an ability to capitalize on, one's own knowledge and thought process as these are applied to some specific task. The general knowledge of the reader guides him or her in monitoring comprehension processes through the selection and implementation of specific strategies to achieve some predetermined goal or purpose for reading. The chief idea involved in metacognition is that learners must actively monitor their use of thinking processes and regulate them according to the purpose for reading.

Presseisen (1984) describes two main dimensions to metacognitive thinking. The first is task-oriented. This dimension has to do with monitoring the actual

performance of a skill. The second dimension is strategic. It relates to how a certain skill is selected above others in a particular circumstance. This includes an awareness of getting the most informative feedback from carrying out a particular strategy. Presseisen claims that in terms of selecting appropriate work strategies, metacognition suggests that the first order of learning is to recognize the particular problem and know what information is needed to resolve it and where to obtain such information. Henle (1966) suggests that recognizing what is understood and to what degree ultimately helps the learner come to terms with the power of his/her own thoughts.

Baker and Brown (1980) divided metacognitive processes into two (not necessarily independent) clusters. The first is concerned with the learner's awareness of any incompatibility between available knowledge and the complexity of the task at hand. The second is concerned with the active monitoring of one's own cognitive processes. According to Baker and Brown, the choice of reading strategies will vary depending on whether the goal is to read for meaning (comprehension) or for remembering (studying). The researchers recognize that reading for remembering involves all the

activities of comprehension and more. Before a review of studies which have focused on metacognition and reading assessment, it is appropriate to discuss the methodology of data collection that was common among them. The interest in the metacognitive aspects of reading has resulted in a debate over the validity of verbal report data. This and most other studies investigating reading processes have included such data.

Verbal Reporting of Reading Processes

The interest in metacognition demanded an exploration of procedures to collect data on thinking processes. How the data are obtained is a major concern in all research studies involving the reading process. One of the earliest studies using a process approach was done by Huey (1912). Subjects were asked to reflect about the method they had used to comprehend vocabulary words both in isolation and in context. The subjects were able to explain that their associations with words in isolation were quite varied, but that they would usually limit the choices to one when they saw the same word in a sentence. One of the most important implications of this study was that readers are aware of and able to verbalize at least part of their reading

processes. Huey's study demonstrated that useful insights about the reading process can be gained by asking subjects about their reading behavior.

This approach to data collection has come to be known as introspection. There are two forms of collecting introspective data. Concurrent verbal reports are collected as the subject is engaged in the experimental task. Retrospective verbal reports are collected after the subject has completed the experimental task.

Concurrent verbal reports have been used as data in research in the field of psychology since early in the 20th century (Pritchard, 1987). During this time period, a movement toward qualitative analysis resulted in what Titchener (1912a) called "systematic experimental introspection". This method of data collection grew out of the belief that some psychological processes were only accessible through self-observation. Titchener and others conducted tightly controlled, laboratory-based studies which only utilized subjects who had gone through extensive training prior to data collection. The subjects were asked to describe their thinking processes as they completed various tasks.

As Titchener was conducting his experiments, the use of introspection was being attacked by some experts in the field of psychology. Pepper (1918) claimed the method was poorly defined and Dodge (1912) that it was inadequately controlled. Watson (1913) did not believe that subjects could observe their own mental process, and claimed that the method was inherently flawed. In response to some of these criticisms, Titchener (1912b) made several modifications to the methodology. He attempted to clarify the conditions under which introspective methods could be employed and how the data could be interpreted. Perhaps the most significant modification was that he distinguished between various types of introspection. This distinction was made on the basis of whether the process and its reporting occurred together (immediate description) or at a different time (retrospection description).

Due in part to the emergence of behaviorism (Bakan, 1954), the use of introspection almost vanished from the research for approximately thirty years. Behaviorists believed that only objective and accessible facts of behavior or activity could be studied. However, the growth of cognitive psychology in the 1950s and 1960s inspired a new interest in introspective methodology as psychologists again became interested in the non-

observable workings of the mind. The advantages and disadvantages of the methodology continue to be examined to this day. One important conclusion of several experts emerged in the mid-eighties. This was that responses of highly trained subjects were artificial and therefore less valid than those of untrained subjects (Meichenbaum, Burland, Gruson, & Cameron, 1985).

Cohen and Hosenfeld (1981) proposed a three dimensional model for researching cognitive processes. This model, attempted to clarify and define the variations of introspection data collection procedures. The model involves the dimensions of activity, time, and content. Activity refers to either thinking-aloud, in which "the subject just lets his or her thoughts flow verbally without trying to control, direct, or observe them" or self-observation, which "can range from largely unanalyzed verbalizations to those that reflect extensive analysis" (1981, p. 286). Time refers to the time elapsed between the activity and the subject's verbalizations of it. A distinction is made between the immediate reporting of the activity (concurrent) and that which does not take place at once (retrospection). Content refers to the need to consider the content of the subject's thoughts since content will vary according to the task. Based on studies which Cohen and Hosenfeld

conducted using this model, they concluded that learners themselves have important insights into their learning processes, and that mentalistic research methods can be valid and reliable mans of investigating those processes.

Despite carefully defined models of collecting introspective data such as Cohen and Hosenfeld's, significant differences of opinion as to the validity and reliability of verbal reports persist. Nisbett and Wilson (1977) claim that all verbal report data are suspect since they believe that people have little or no access to their higher order cognitive processes. However, several other prominent researchers believe that verbal reports are useful under certain conditions. Kanfer (1968) maintains that introspective data are useful in exploratory studies as a means of identifying hypotheses which could be examined in subsequent research. Others state that introspective data is useful as one of several data bases providing convergent validity for a given research question (Kail & Bisanz, 1982; Ross, 1983; Sternberg, 1985). These researchers suggest limitations regarding the conditions of the use of verbal report data.

Other researchers maintain that verbal reports, when they are elicited with care and interpreted with

full understanding of the circumstances under which they were obtained, are valuable and thoroughly reliable source of information about cognitive processes (Afflerbach & Johnston, 1984; Ericsson and Simon, 1980; Meichenbaum et al., 1985; Smith & Miller, 1978; White, 1980).

The use of self-report data for understanding reading processes has gained in popularity in research on reading (Afflerbach & Johnston, 1984). When used in reading research, concurrent verbal reporting refers to the process whereby subjects are asked to describe what they are thinking and doing as they are reading the text. The use of concurrent verbal reporting has been criticized for interfering with the normal processes of reading since subjects must not only attend to the task of reading, but also must report their thoughts surrounding the text (Nisbett & Wilson, 1977; Garner, 1982).

In studies using retrospective verbal reports for data collection, subjects are asked to describe what they were thinking and doing in relation to the text after the complete text has been read. This system has been criticized because it relies on the subject's ability to accurately remember and report his reading behavior (Afflerbach & Johnston, 1984).

Most research studies in reading processes have used either concurrent or retrospective data collection techniques, but not both. Because each of the two techniques have their own set of strengths and limitations, it was decided to include both in this study. Furthermore, the subjects in this study completed each task in an observation-only setting. During these sessions, subjects were not asked to report any information. Eye movement data have been suggested as potentially important indicators (Afflerbach & Johnston, 1984; Carpenter & Just, 1981; McConkie, Hogaboam, Wolverton, Zola, & Lucas, 1979). By comparing the eye and hand movements of subjects as they were only being observed to the descriptions of what the subjects reported during concurrent and retrospective verbal reports, the validity of the verbal reports will be expanded.

In studies of metacognition and reading, researchers have compared reading processes of various types of readers and under a variety of conditions. Some examples are comparisons of older versus younger readers (Garner & Reis, 1981), more proficient versus remedial readers (Garner & Reis, 1981; Olshavsky, 1976-1977), readers dealing with high-interest versus low-interest texts (Olshavsky, 1976-1977), expert versus

novice readers (Lundeberg, 1987; Johnson, 1984); and readers reading culturally familiar versus culturally unfamiliar texts (Pritchard, 1987). Only a handful of studies have examined reading processes as they relate directly to reading assessment. Before these studies are reviewed, a review of the major issues related to the validity of reading comprehension assessment will be presented.

Reading Comprehension Test Validity

Perhaps there has never been a more controversial and seemingly endless debate in the field of reading than that concerning reading assessment. Concepts of both literacy and its assessment have changed slowly over time. In the Middle Ages when Latin was the lingua franca of religion and law, a litteratus was one learned in Latin (Clanchy, 1983). This definition contrasts sharply with the early history of literacy and its assessment in American History. In 1840 the Census Bureau began to include a literacy question in its household surveys without any criterion as to what constitutes literacy. Respondents were asked if they could read and write and a simple yes or no was recorded (Venezky, 1986). Formal measures of reading and writing ability had yet to be developed.

Around the turn of the century, the testing of reading and writing in American schools became popular. Tests were used as a means to evaluate the effectiveness of schools (Farr & Carey, 1986). By 1917, eighteen U.S. cities had research organizations within their school systems constructing, administering, and interpreting tests (Levine, 1976).

Today, the assessment of reading and writing is a complicated and controversial topic. Multi-million dollar businesses create tests which are purchased by school corporations and taken by virtually every student in American schools, sometimes several times each year. Results from these tests are used by school corporations in accountability studies, teacher evaluations, student evaluations, student placement and programming, program evaluation, admission to institutes of higher learning and even sometimes admission to kindergarten (Hoopfer & Hunsberger, 1986).

There are numerous allegations surrounding reading comprehension tests. Critics of standardized tests have cited lack of objectivity, test bias, over-dependence on reading, over-dependence on statistical power, lack of breadth and depth of content covered, penalties for critical thinkers, penalties for careless bookkeeping, ambiguity of text and questions, reification of test

scores, control of the curriculum by test constructors, and lack of diagnostic value and information gain (e.g. Hoffman, 1962; Houts, 1977; Strenio; 1981). Other concerns include the assertions that the passages used in the test are out of context, the imposition of adult reality on child perception, the limits to the certainty of assessment, the incomplete use of rules during the testing and the assumption that informational links are the same for all children (MacKay, 1974; Roth, 1974; Mehan & Wood, 1975; Heap, 1980).

Valencia and Pearson (1987) claim that reading assessment has not kept pace with advances in reading theory, research, or practice. They have argued that over the last few decades the field has gained a better understanding of a strategic view of the process of reading (e.g., Collins, Brown, & Larkin, 1980; Pearson & Spiro, 1980). Valencia and Pearson claim that progress toward expert reading is guided by the increasing sensitivity of readers to issues of <u>how</u>, <u>when</u>, and <u>why</u> to use resources such as text, prior knowledge, and environmental clues, as well as the ability to be flexible in the attempt to make meaning of print.

At the core of these and other controversies over the validity of current forms of reading assessment is

whether reading comprehension is a process or a product. Many critics of reading tests claim that most current approaches to the assessment of reading comprehension embody an assumption that reading comprehension is a product of a reader's interaction with a text. The processes involved in comprehension are less emphasized than the final product (Johnston, 1983). Royer and Cunningham (1978) contend that comprehension processes and memory processes are inextricably intertwined. They assert that a comprehended message will be retained in memory better than an uncomprehended message. Bower (1978) believes that superior memory seems to be an incidental byproduct of fully understanding a text. While these assertions may seem to be only common sense, they have important implications for reading tests. If students comprehend a text, their mental processes have been involved with it. Therefore, they will remember the text better than someone who hasn't comprehended the text. As a result, they will gain higher scores on reading comprehension tests.

While the debate over process versus product may never be fully resolved, the two do seem to be interrelated. Reading comprehension test scores are the product of reading. The fact that the score (product) of these tests are emphasized more than the processes

involved in getting them is more a reflection on the curriculum and the people interpreting the test than the test itself. Virtually all assessments of reading are indirect, even those that claim to directly assess reading processes. Even self-report data are products. We cannot actually see the processes involved, we can only infer from the verbal reports what the subject is thinking. Nor can we get an exact measure of how much has been comprehended, we can only get an estimate of the amount that has been comprehended. Therefore, any test of reading, including verbal reports given by students while reading, are indirect, and thus products of the reading process.

The fact that educators have access only to the products of actual reading processes pose serious problems for the validity of assessing reading comprehension. Validity is the most important aspect of any test. Construct validity is of particular importance. Questions of construct validity are concerned with whether a reading test measures the true behavior of reading. Farr (1986) states that "the manuals of most standardized tests make very explicit the fact that the test will not provide information about a pupil's reading processes, but only information about the product of reading". However, he continues by

saying that "...one could argue that the product-or score- isn't valid if a pupil doesn't use the actual processes of reading in determining the answers." The validity question that surrounds the tests then seems to be whether or not taking the test appears to change the reading processes involved in comprehension. Therefore, this study has been designed around that question.

Research that focuses on the metacognitive aspects of reading while taking a reading test comprise only a small portion of the literature. Only recently have researchers begun to take an interest in what thinking processes the reader may use during reading that may, or may not, facilitate remembering of text.

Wingenbach (1984) examined the comprehension processes employed by 100 gifted readers in grades four through seven to identify the various metacognitive strategies they employed as they read the Iowa Tests of Basic Skills (ITBS). After taking the test, all subjects completed a questionnaire to determine metacognitive awareness of strategy selection and use. The questionnaire was designed to ask students about strategy usage. Metacognitive awareness in the form of selection, use, and evaluation of strategies was the focus of the questionnaire. Students were asked to

evaluate frequency of use of specific behaviors under conditions set out in the question.

Based on grade equivalence results of the ITBS, the top five students in each grade level (a total of twenty subjects) were selected to participate in protocol analysis of the comprehension process and an interview. These students were given two short passages with related questions from the ITBS and were asked to stop at indicated points to describe the text and what they were thinking.

The same twenty students who participated in protocol analysis also participated in an interview. The interview questions were developed to reflect those reading strategies indicated as used frequently by responses to the metacognition questionnaire. Following the protocol analysis session, each student was interviewed concerning his or her strategy use.

Patterns/behaviors as described by the students were classified into categories of strategies already established by Olshavsky (1976-77). One new strategy category, the use of imagery, was established as behaviors involving imagery were described by the subjects.

Findings revealed that subjects perceived their use of the following strategies which Wingenbach describes as follows:

Reading Strategies: Three strategy groups. Strategies may occur singularly or be used simultaneously.

I. Word related strategies: pertains to word meaning within passage.

context - use of context to define a word or use of contextual information to define a word.

synonym substitution - use of a synonym in place of a word used in the text. Synonym substitution is used to check the meaning of a word in the text or to "translate" the text into a more meaningful statement for the reader.

failure to understand a word - the subject states that the word is unknown. The use of other strategies may then follow to identify the word.

II. Context related strategies: pertains to understanding statements or ideas as presented within the text.

rereading - returned to the text for reading again for one or several reasons: did not understand the first reading, need clarification, an idea needs reinforcement, etc.

inference - interpretation of the text based on information within the text.

failure to understand a clause - subject states an inability to understand a clause or sentence or paragraph. Other strategies such as rereading may then be applied.

use of information about the story - information from the immediate text is related to the overall story or used to understand new information in the story.

III. Meaning related strategies: employed when the subject is attempting to gain an understanding of a larger concept than the literal statement of text. These strategies usually require the interaction of several knowledge sources.

personal identification - the subject relates what is in the context to personal experience(s) and usually elaborates the contextual statement in some way emotionally, visually, or informationally.

addition of information - the subject adds personal knowledge to the information provided in the text. The addition usually elaborates the context and enriches comprehension.

hypothesis/anticipation - subject predicts or indicates awareness of what will follow based on words in context, subject's own experience and/or prior knowledge.

imagery - subject visualizes the meaning of the text. This may involve addition of information, comparisons, etc. Imagery may be used for clarification or as a check for understanding.

Wingenbach states that the protocol analysis and interview components supported the results of the traditional research procedures. During the protocol analysis session specific strategies were used and discussed by participants. These same strategies became the focus for the interviews. The researcher states that students found that during the interview it was easy to discuss strategies such as rereading and the use

of context. She claims that perhaps the previous experience with the questionnaire and the protocol analysis sessions made the strategies and their use more explicit and thus easier to discuss. Although this piece of research made a good attempt to validate the strategies through more than one interview with each subject, two limitations are evident. First, the researcher did have a predetermined set of strategies prior to conducting the research, that being Olshavsky's (1977-77). Secondly, all of the subjects may also have had a predetermined set of strategies in mind prior to the protocol analysis since they had all previously completed the questionnaire that was based on Olshavsky's set of strategies. These prior experiences with what the subject's may have perceived as the researchers idea of strategies may have influenced what they said in subsequent interviews. This is one of the major criticisms of using "trained" subjects.

Alvermann and Ratekin (1982) conducted a study with ninety-eight seventh and eighth grade "average" subjects to determine among other things whether or not strategies that students reported they used during reading in order to complete a multiple-choice test differ in type or incidence from those which they reported they used during reading to complete an essay

test. A multiple-choice test and an essay test were constructed to test the comprehension of two folk tales. Each multiple choice test consisted of five literal and five inferential questions. The essay test consisted of two written sets of directions to the student. The first set directed the student to write the moral or lesson that the folk tale taught. The second set required a brief description of how the identified moral might apply to the student's life.

Each subject was assigned to either multiple-choice or essay test groups. Prior to reading, the students were told to read the folk tale and to prepare for the test appropriate to their group. A standardized retrospective interview was given to determine what strategies the subjects remembered using as he or she read the folk tale. Again, Olshavsky's (1976-77) method of identifying and categorizing strategies was employed. An analysis of strategies which students reported they used during reading resulted in identification of the following seven categories by the researchers. An example of the protocol was used by the researchers to illustrate the meaning of each category.

1. <u>Reread</u>

> "I read it over - each paragraph twice - until I remembered what it said."
> "I remember it by going over the story two times."

2. Read carefully/slowly

"I read the story very carefully and I thought I won't remember it but it always comes back."

3. Read for details

"Read it so I could remember in detail what the story was about by remembering some of the words."

4. Read for main ideas

"Tried to remember the main ideas."

"I looked for the most important parts of the story. For example, the man that collected the carvings of animals did not planned (sic) on being fooled. The second carver planned his idea out very carefully because he wanted to get the bag of gold."

5. Personal Identification

"I thought of how I would feel if I was the dragon."

6. Imaging

"The way I remember a story is I put pictures in my head as the story goes along."

7. No specific strategy

"When I read I remember it in my head."

"I just read it good and then when I answer questions it comes back to me."

Examples of multiple strategy use:

"I reread the story and read it slowly."

"I read it once and then I read it over again to make sure I didn't miss any details."

Fifty-five students reported using only one strategy while thirty reported using two or more

strategies. Thirteen students were unable to recall any specific strategy. Interrater agreement between two independent judges who classified each of the strategy statements was .93.

Fisher's exact probability test showed that only the strategy "rereading" yielded a significant difference ($p < .05$) between students in the two task groups. Students who read and studied for an essay test "reread" more frequently than students who read the passage with a multiple-choice test in mind. Students who read to complete an essay test reported using multiple strategies nearly twice as often as students who read for a multiple choice test. The proportion was .40 for the essay group compared to .22 for the multiple choice group.

The researchers only elaborate on the statistically significant differences in the text of their report. However, an examination of a chart representing the frequency of reported strategies shows that students read for details twice as often in the multiple choice test as they did in the essay test. There were four reports of imaging in the essay test compared to one in the multiple choice test. Subjects made a personal connection with the text seven times in the multiple choice test compared to three in the essay test. With

such small numbers, it is difficult to get any statistically significant difference. While these differences are not statistically significant, they do illustrate differences that may be significant from a pragmatic point of view.

Alvermann and Ratekin state that the fact that "reading carefully/slowly" and "rereading" were the strategies of choice is somewhat disturbing because of their generally passive nature. They also state that the results suggest that seventh and eighth grade average readers may only have a limited awareness of the entire range of strategic activities available, or the alternative hypothesis, that these students revealed the nature of what they found inherently useful from past reading instruction. They conclude that the fact that they did not report using other strategies may have been due more to a failure to recognize the need for strategic intervention than to either limited awareness or prior instruction.

This study was one of the few in the literature that looked specifically at how reading processes may vary due to the nature of a testing task. However, the study has major limitations which are not cited by the researchers. First, the use of only retrospective data severely limits the conclusions. When retrospection is

used alone, the chances that the subjects forgot the mental processes they employed while reading are greatly enhanced. The researchers claim that the subjects' failure to report other strategies may have been due to a failure to recognize the need for strategic intervention than to limited awareness. They do not recognize the possibility that only these strategies were reported because subjects may have forgotten using other strategies. Secondly, the differences which were found could be accounted to individual or group differences rather than task-related differences, which is what the researchers had intended to explore. There is little information in the report that the two groups were equivalent. It is reported only that they all had obtained stanine scores of 4, 5, and 6 on the reading subtest of the Iowa Tests of Basic Skills five months prior to data collection. Finally, there is no evidence in the report that the categories were validated either through the literature or through independent readers. They were simply put into Olshavsky's preexisting categories.

Afflerbach & Johnston (1984) and Fareed (1971) are among those who advise not to force protocol responses into preexisting categories, which both studies did. By forcing the protocols into preexisting categories,

the researchers already had a mind set on what strategies exist. They may therefore have overlooked responses that were subtly, yet pragmatically significant.

In summary, these two studies have utilized introspective approaches to gain insights into various aspects of the reading process under testing conditions. Both of the studies had serious flaws in the data collection procedure and in the analysis of the data. In spite of the numerous problems with these studies, they do provide a basis upon which future studies can be designed. Studies need to be conducted that do not fit responses into existing categories. The categories should emerge from the data and be validated by independent readers. Subjects should not be exposed to a predetermined set of strategies prior to data collection. This may lead them to reporting what they think the researcher wants to hear. More than one data collection technique should be employed. The strengths and weaknesses of both concurrent and retrospective data collection needs to be balanced. Differences in strategy use by task should be examined within individuals, not groups. There are too many factors that can contribute to differences across groups.

Summary

This chapter has established a theoretical and methodological framework for an investigation of the reading comprehension processes used by readers as they engage in taking reading comprehension tests. The next chapter presents a description of and rationale for the specific procedures used to conduct this investigation.

CHAPTER III

METHODOLOGY

This study was designed to describe and compare the reading comprehension processes of students as they engaged in a multiple choice test, a cloze test, a written retelling, and a nonassessed reading activity. This chapter describes in detail the methodology used to investigate this research question.

Pilot Study

A pilot study was conducted to assess the validity of the procedures and materials developed for this investigation. The use of observation, concurrent and retrospective verbal reporting seemed appropriate measures of data collection given the nature of the study. Therefore, the pilot study was needed to determine if sixth-grade proficient readers were capable of providing the necessary verbal protocols that would justify the use of concurrent and retrospective verbal reports.

Three sixth grade proficient readers were selected to participate in the pilot study. These students attended a school which was approximately five miles from the schools that were to be used in the actual study. Therefore, it was unlikely that the subjects in the pilot study would discuss the study procedures with the experimental subjects.

Based on the pilot study, two things were modified from the original proposal. First, it was determined that more information was reported by subjects as they gained experience in giving verbal reports. Therefore, it was determined that practice passages should be used in the actual experiment. Secondly, it was found that the more difficult the text, the more information was reported. This was probably the case because when comprehension is relatively effortless, the processes are too automatic to be reported by proficient readers (Johnson, 1984; Polanyi, 1973). Therefore, some of the sixth grade passages were replaced with passages of seventh and eighth grade readability as reported in the MAT. No other significant changes were made after the pilot study.

The pilot study demonstrated that sixth grade proficient readers understood the directions for both retrospective and introspective reports and were able

verbalize the processes used while reading. Based on the pilot tryouts and prior research, it was concluded that the procedures and subjects were appropriate to this study.

Subjects

The nature of this study dictated a small sample size. The purpose was to examine individual differences in comprehension processes across the four tasks. A total of one hundred and eight interviews were conducted (twelve interviews with each of nine subjects). Had this number of interviews been conducted with different subjects, too many variables could have been the cause for any differences in the reading processes observed. Therefore, nine subject were interviewed twelve times. Each subject completed three multiple choice tests, three cloze tests, three written retellings, and read three nonassessed passages.

The small sample for this study does limit the generalizability of the results. Afflerbach and Johnston (1983) report that most verbal report studies have used relatively small numbers of subjects. The amount of time needed to transcribe and analyze protocol contents may be a concern for researchers considering using verbal reports. Johnston and Afflerbach (1983)

found that their subjects often supplied 15-20 typewritten pages of their processes for a single, 5-paragraph passage. Because of this and the fact that the purpose of this study was to describe in-depth the individual reading processes on the four given tasks, it was felt that the small sample would permit a more thorough understanding as to how a reader's processes alter under various testing conditions. This in-depth analysis provides information that may generate hypotheses for larger scale studies.

Sixth graders were an appropriate grade level for this study for three reasons. Students at this age are capable of reading the passages and are able to verbalize their comprehension processes. In addition, while sixth graders were likely to be more experienced in taking multiple choice tests than cloze tests or retellings, they are not as sophisticated test takers as an older population would have been.

Because poor readers lack understanding of the strategic nature of reading (Canney & Winograd, 1979; Markman, 1977, 1979; Paris & Lindauer, 1976), only proficient readers were used. Proficiency was determined by teacher identification instead of standardized reading test scores since the focus of the study is directly related to reading tests. The

subjects were selected in the following manner. The researcher met individually with two sixth grade teachers. At these meetings, each teacher was asked to read and follow the directions below which were adapted from Hopkins, George, and Williams (1985) and Pritchard (1987):

> I am conducting a research study which will investigate the reading comprehension processes of sixth grade students as they engage in different forms of reading assessment tasks. I would like for you to identify proficient sixth grade readers from your class. Try not to be influenced by the pupil's effort, interest, attitudes, intelligence, behavior, or attendance. Rate only the pupil's current achievement in reading.
>
> Try to think of one student in your class who you feel can read proficiently at the sixth grade level. Once you have done this, compare the other students to this one. Identify others that are equal to or above your model in terms of their reading ability. Do not include anyone who cannot adequately read sixth grade material.
>
> All of your ratings will be confidential. No teacher or student will be identified in my report. Thank you for your cooperation.

No emphasis was given to the socioeconomic status of the subjects. However, according to the superintendent of the school corporation, the schools in which these subjects were enrolled represented a wide range of socioeconomic levels.

In order to determine if there were differences in reading processes according to sex, a stratified random

sample was taken in order to obtain five girls and four boys for the population.

After the necessary formal permission for research was returned by the school corporation, parental permission forms were sent home and returned by a total of fourteen students. Of these, a stratified random sample was used to get the following population; five girls and four boys, five subjects from one school and the remaining four from the other.

Materials

Twelve expository passages from the 1984 edition of the Metropolitan Achievement Test (MAT) were selected. Passages from the MAT were selected because they are popularly used tests and they are representative of the standardized reading comprehension multiple choice tests currently being used throughout the country. In addition, the passages in the MAT are approximately the same length and the basis for determining the difficulty of the passages is consistent across passages. The difficulty of the passages in the MAT are determined in the following manner as described in the MAT 6 Reading Diagnostic Tests Teacher's Manual for Interpreting Tests (1987):

> The difficulty of the passages has been controlled so that, in general, they match the reading difficulty of the basal readers; that is, the passages which have been designated as Grade 4 are similar in reading level to passages found in the Grade 4 basal readers. This was done by controlling vocabulary, sentence and passage length, difficulty of concepts, and interest level of the topics. For passages designated Primer through Grade 7-8, the grade level determinations were based on a vocabulary control formula using Albert J. Harris and Milton D. Jacobson, Basic Reading Vocabularies, New York: Macmillian, 1982.

Expository passages were used for two reasons. First, it was felt that the passages should be as similar to each other as possible. Therefore, it was felt that either exclusively expository or exclusively narrative passages should be selected. The majority of passages on the MAT at the sixth, seventh, and eighth grade level are expository. Hence, expository passages were selected. The passages were determined to be expository on the basis that they had no evidence of following a story grammar. All of the passages presented information on a certain topic or person.

The passages were rotated within the four tasks to control for background knowledge of the content, passage length and reading difficulty. The following chart represents the assigned passage number, the topic of the passage, the number of words in the passage, and the readability level assigned to it by the MAT:

TABLE 1

Nature of Passages

Passage	Topic	# Words	Readability Level
1	Maple Syrup	145	6
2	Sports Figure	155	6
3	AerOcycles	235	7-8
4	Weeds	180	6
5	Gazebos	174	6
6	Monopoly	211	7-8
7	Sports Figure	208	7-8
8	Turtles	173	6
9	Plywood	183	6
10	Sherlock Holmes	213	7-8
11	Icebergs	176	6
12	Oil Rigs	226	7-8

Average passage length = 190 words
Average sentence length = 13 words
Range in passage length = 145-235 words

Data Collection Procedures

Oral reporting of reading strategies has been widely defended as a valid means of gaining information about the reading process (Ericsson & Simon, 1980; White, 1980; Afflerbach & Johnston, 1984; Baker & Brown, 1984; Hynd & Alvermann, 1984). There are two commonly used methods of gathering this type of data. The first is concurrent verbal reporting. During these interviews, the subjects are asked to describe what they

are thinking and doing as they complete the task. The second method is retrospective verbal reporting. In retrospective interviews, the subjects are asked to report what they were doing and thinking as they read after the task has been completed.

The rationale for using both concurrent and retrospective verbal reporting as data collection procedures is because the literature is divided as to what differences, if any, there are in the nature of the information that is derived from each (Cohen, 1986). Concurrent reporting has been criticized for interfering in normal task process, thereby yielding invalid information (Nisbett & Wilson, 1977; Garner, 1982). Retrospective verbal reporting has been criticized for yielding incomplete results due to the subject's forgetting details of his or her mental processing (Afflerbach & Johnston, 1984). Both have been criticized for leading students to tell the researcher what the subject believes the researcher wants to hear. The inclusion of both interview methods was an attempt to balance the strengths and criticisms of each method. Also, the observation of the subjects as they completed the tasks without using either types of verbal reports served as an additional level of credibility to the data collection procedures.

The purpose of having the subjects read one passage without any formal assessment was to have as a comparison one set of reported processes where subjects were not anticipating a formal assessment of their comprehension. It is recognized that the context of the situation under which this information was gathered made it an "unnatural" reading act. However, these nonassessed reading passages were read under the same physical conditions as the tests were taken (away from the classroom with the researcher present). Since the conditions were the same, any differences in processes reported between the assessed tasks and the nonassessed tasks are likely to be due to the task, not the manner in which the data were collected.

During the first four interviews, as each subject took one multiple choice test, one cloze test, engaged in a written retelling, and read a nonassessed passage, the subjects were only observed, no verbal responses were collected from these activities. The researcher observed the eye and hand movements of the subjects as they completed the four tasks and made notes of her observations. During these sessions, the researcher was able to determine in a general way how the subjects went about completing the task. For example, during the cloze test, the researcher was able to determine that

the subject's eyes scanned over each sentence several times before filling in a blank. This served as a baseline for interpreting the verbal protocols of the cloze testing. During both types of verbal reporting, all subjects reported rereading a great deal during the cloze test. As stated earlier, this had been observed in the observation-only sessions, as the subjects eyes had appeared to be going back and forth over the same section of text. The observations of each subject were extremely similar across tasks. Therefore, a summary of the observation-only interviews was written by task and is presented in Chapter 4.

After the initial observation-only sessions, subjects were given one practice passage before each type of verbal reporting data was collected. As the subjects reported concurrently, they described as they completed each task what they were thinking and doing. As the subjects reported retrospectively, they recalled what they were thinking and doing after they had completed the task.

In an attempt to expand the meaning of the verbal protocol data and to gather information about how each subject viewed each type of task, an "Exit Interview" was given. During the final interview with each

subject, the subjects were asked to think back over each type of task and answer the following questions:

EXIT INTERVIEW

1. How did you go about completing the multiple choice test?

2. How did you go about completing the cloze test?

3. How did you go about completing the written retelling?

4. Think about the passage you read when you didn't take a test. How did you read it? Which test passage did you read the most like you read that passage?

5. Which test was the hardest? Why?

6. Which test was the easiest? Why?

7. If you were a teacher and you wanted to really see how well your students could read, which test would you give them? Why?

Subjects were excused one at a time from class by their teacher to participate in the study. Every effort was made to ensure that the meetings with each individual were as comparable as possible. In one school, interviews were held in a room off the library. In the other school, the interviews were held in the speech and hearing room. Most of the time both rooms were relatively quiet. However, as is the case in most schools, occasionally noise from the hall could be heard.

The data collection took place over a period of twelve weeks. It had been intended that each session would be conducted in the morning. However, as the school year was coming to a close, five of the final interviews at one school had to be conducted in the afternoon because of scheduling problems. Since this was so late in the data collection, the subjects were already very familiar with both the researcher and the procedure, and the subjects did not appear to be affected by the change.

The average length of the interviews was twenty minutes. Due to the fact that subjects could only miss their regular classroom studies during times determined by their teachers, it was impossible to interview all subjects on any given day. Usually only two interviews could be conducted in one day. Therefore, a rotation system was used. For example, if subjects 1 and 2 were interviewed one day, subjects 3 and 4 would be interviewed during the next visit to that school. Only one interview was conducted for any subject per day.

The data collection period overlapped at each school. They were conducted over a ten week period on each Monday, Wednesday and Friday at one school. The interviews at the second school were conducted over a twelve week period on Tuesday and Thursday. Because

only two days per week were spent at this school, two additional weeks were needed to finish the data collection after the completion of the data collection at the other school.

The chart on the following two pages provides an overview of the data collection procedure. This chart reports the order in which the following were given to each subject; the type of data collection, the passage number, and the type of task. In order to clarify how this worked, a detailed account outlining the procedure as it applied to one subject is warranted. Therefore, a detailed account of the procedure as it applied to Subject #1 follows the chart.

TABLE 2

DATA COLLECTION PROCEDURE

Subject #1	Subject #2	Subject #3
Observation	Observation	Observation
1*MC	12 CT	11 WR
2 CT	1 WR	12 NA
3 WR	2 NA	1 MC
4 NA	3 MC	2 CT
Retrospect	Concurrent	Retrospect
5 CT	4 WR	3 NA
6 WR	5 NA	4 MC
7 NA	6 MC	5 CT
8 MC	7 CT	6 WR
Concurrent	Retrospect	Concurrent
9 WR	8 NA	7 MC
10 NA	9 MC	8 CT
11 MC	10 CT	9 WR
12 CT	11 WR	10 NA

Subject #4	Subject #5	Subject #6
Observation	Observation	Observation
10 NA	9 MC	8 CT
11 MC	10 CT	9 WR
12 CT	11 WR	10 NA
1 WR	12 NA	11 MC
Concurrent	Retrospect	Concurrent
2 MC	1 CT	12 WR
3 CT	2 WR	1 NA
4 WR	3 NA	2 MC
5 NA	4 MC	3 CT
Retrospect	Concurrent	Retrospect
6 CT	5 WR	4 NA
7 WR	6 NA	5 MC
8 NA	7 MC	6 CT
9 MC	8 CT	7 WR

Subject #7	Subject #8	Subject #9
Observation	Observation	Observation
7 WR	6 NA	5 MC
8 NA	7 MC	6 CT
9 MC	8 CT	7 WR
10 CT	9 WR	8 NA
Retrospect	Concurrent	Retrospect
11 NA	10 MC	9 CT
12 MC	11 CT	10 WR
1 CT	12 WR	11 NA
2 WR	1 NA	12 MC
Concurrent	Retrospect	Concurrent
3 MC	2 CT	1 WR
4 CT	3 WR	2 NA
5 WR	4 NA	3 MC
6 NA	5 MC	4 CT

* Passage number

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Non-assessed Passage

Figure 1

OUTLINE OF PROCEDURES USED WITH

SUBJECT #1

Interview #1

The researcher spends approximately five minutes talking to the subject in an effort to establish rapport. The researcher explains that she is interested in learning about how sixth grade students read and take reading tests. The subject is asked if she has any questions.

Passage #1 is presented to the subject in the form of a multiple choice test. The researcher explains that once the subject has started a task, no assistance would be given by the researcher. The subject is given the following instructions orally by the researcher:

"Read the article and choose the best possible answer to each question. Try to take the test exactly as you would if you were taking it in your classroom for a grade."

Subject is observed as she completes the task.

Interview #2

Passage #2 is presented to the subject in the form of a cloze test. The subject is given the following instructions orally by the researcher:

"As you read the article, fill in the blanks with the word that best completes the sentence. Try to take the test exactly as you would if you were taking it in your classroom for a grade.

Subject is observed as she completes the task.

Interview #3

Passage #3 is presented to the subject. The subject is given the following instructions orally by the researcher:

"Read the article. When you have finished I am going to take the article away from you. Then I will instruct you to write down everything you can remember about the article. Try to do this just as you would if you were in your classroom doing this for a test.

Subject is observed as she completes the task.

Interview #4

Passage #4 is presented to the subject. The subject is given the following instructions orally by the researcher:

"Pretend that you are in the waiting room of your doctor's office. You are very bored, so you decide to read this article. Try to read this article just as you would under those conditions."

Subject is observed as she completes the task.

Interview #5

A non-assessed practice passage is presented to the subject. The use of retrospection is explained to the subject in the following way:

"As you are reading this article, try to think about what you are thinking and doing in order to understand what it says. When you have finished reading it, I am going to ask you to describe what you were thinking and doing."

After subject has completed the practice task, passage #5 is presented to the subject in the form of a cloze test. Subject is again given the instructions for taking a cloze test, and told to try to remember what she was thinking and doing as she takes the test so that she can describe it after the test has been completed.

After the subject has completed the task, she reports what she had been thinking and doing.

Interview #6

Passage #6 is presented to the subject. Subject is read instructions for written retellings and retrospection.

After the subject has completed the task, she reports what she had been thinking and doing.

Interview #7

Passage #7 is presented to the subject. Subject is read instructions for the non-assessed task and retrospection.

After the subject has completed the task, she reports what she had been thinking and doing.

Interview #8

Passage #8 is presented to the subject in the form of a multiple choice test. Subject is read instructions for multiple choice tests and retrospection.

After the subject has completed the task, she reports what she had been thinking and doing.

Interview #9

A non-assessed practice passage is presented to the subject. The use of concurrent verbal reporting is explained to the subject in the following way:

"I want you to talk to me as you read this passage. Please stop at the end of each sentence, or whenever you can think of something to say, and tell me what you are thinking and doing. You do not need to read the article aloud."

After subject has completed the practice task, passage #9 is presented. Subject is given instructions for a written retelling, and told to tell the researcher what she is thinking and doing as she completes the task.

As subject completes the task, she reports to the researcher what she is thinking and doing.

Interview #10

Passage #10 is presented to the subject. Subject is read instructions for nonassessed reading task and concurrent verbal reporting.

As subject completes the task, she reports to the researcher what she is thinking and doing.

Interview #11

Passage #11 is presented to the subject in the form of a multiple choice test. Subject is read instructions for multiple choice tests and concurrent verbal reporting.

As subject completes the task, she reports to the researcher what she is thinking and doing.

Interview #12

Passage #12 is presented to the subject in the form of a cloze test. Subject is read instructions for cloze tests and concurrent verbal reporting.

As subject completes the task, she reports to the researcher what she is thinking and doing.

Subject is given "Exit Interview" to review thoughts and feelings about each type of task.

During the first session, approximately five to ten minutes was spent talking to each subject in an effort to establish rapport. During this time, the subject was asked general questions such as how long they have gone to their school and how many brothers and sisters they have. Following this period of discussion, the subjects were told that the researcher was a former elementary

school teacher and was now a doctoral student in reading education. It was explained that this research would be used in her doctoral dissertation. The only explanation for the research given to the subjects was that the researcher was interested in learning more about how sixth graders read and take reading tests.

Prior to the beginning of each session, the subjects were reminded that they would not be assisted or interrupted once they started the task. The following instructions were given to each subject for each task, regardless of whether or not observation, concurrent or retrospective verbal reporting was being used:

Figure 2

DIRECTIONS FOR EACH TASK

Multiple Choice

> Read the article and choose the best possible answer to each question. Try to take the test exactly as you would if you were taking it in your classroom for a grade.

Cloze Test

> As you read the article, fill in the blanks with the word that best completes the sentence. Try to take the test exactly as you would if you were taking it in your classroom for a grade.

Written Retelling

> Read the story. When you have finished I am going to take the article away from you and ask you to write down everything you can remember. Try to do this just as if you would if you were in your classroom doing this for a test.

Nonassessed Reading

> Pretend that you are in the waiting room of your doctor's office. You are very bored, so you decide to read this article. Try to read this article just as you would if this were the case.

Observation-Only Interviews

During the observation sessions, the subjects were given the above directions for each task. They were told that no assistance would be given once they started the task. They were asked to complete the tests just as they would if they were taking them in class for a grade. They were told to read the nonassessed reading activity as if they were sitting in a doctor's office. No further instructions were given. The researcher made notes of her observations.

Retrospective Interviews

During the retrospection interviews, the subjects were given the instructions for each task as described above. In addition, they were told the following:

> As you are completing this (multiple choice test, cloze test, written retelling, or reading activity), try to remember what you were thinking and doing to complete the task. When you have finished with the task, I am going to ask you to describe what you were thinking and doing.

Once the task was completed, the researcher asked the following questions:

1. What were you thinking and doing as you read the article?
2. (In the assessed tasks) How did you come up with your answers?

When the subjects stopped talking, they were asked, "Can you think of anything else?" as a prompt. This was done until they answered "no". At that time, the interview was considered complete.

Concurrent Interviews

During the concurrent interviews, the subjects were given the same instructions for each task as described above. In addition, they were given the following instructions:

> I want you to talk to me as you complete this (multiple choice test, written retelling, cloze test, or reading activity). Please stop at the

end of each sentence, or whenever you can think of something to say, and tell me what you are thinking and doing. You do not need to read the article aloud.

In addition, the following directions were given for each assessed task:

Multiple Choice:

As you select the answers, describe to me how you determined that they were correct.

Cloze Test:

As you fill in the blanks, describe to me how you decided that that word fit there.

Written Retelling:

As you write the retelling, describe to me how you decided what should be written.

If subjects were silent for more than approximately 30- 45 seconds, they were reminded to tell the researcher what they were thinking and doing. If it was unclear to the researcher what the subjects were describing, the subjects were asked to explain more clearly.

The data for this study consisted of the researcher's observations of each subject's eye and hand movements, as well as the readers' verbal protocols describing what they thought and did as they engaged in each task. The concurrent and retrospective interviews

were tape-recorded and verbatim transcriptions were taped for later analysis.

All data were collected by the researcher. The researcher, a former elementary teacher of six years, was experienced with sixth grade students. This experience was advantageous because it helped her to establish rapport with the students and put them at as much ease as possible with the research tasks.

The verbal protocols were transcribed by the researcher within forty-eight hours after each interview. A random sample of approximately thirty percent of the verbal protocols were selected for verification of the accuracy of the transcriptions. This was done by a person other than the researcher. Approximately twenty word omissions or additions were found and corrected after consulting with the researcher. In five instances, the person listening to the tapes disagreed with what the transcripts reported the subjects as saying. These instances occurred when it was hard to understand the subject's verbalizations. In these cases, the original versions of the transcripts were used since they had been transcribed by the researcher within a relatively short time period after the interviews.

The concurrent and retrospective verbal reports were transcribed by idea units and numbered. This was done so that the verbal reports could later be categorized by reading processes. Idea units were defined as a group of related words that contains both a subject (stated or understood) and a verb phrase which, with its modifiers, forms a single idea. This definition was based on one created by Pritchard (1987). This definition of idea units was used because, as Pritchard points out, the more common definition of idea units, that is "functional boundaries based on pausal acceptability" (Johnson, 1970), is based on surface structure conventions. These conventions ignore the semantic content of the text and therefore was a less appropriate basis for analysis.

A person other than the researcher independently read each protocol to check for the accuracy of the identification of the independent clauses. There was 92% agreement. Those for which there was disagreement were discussed until an agreement was reached.

A sample of the raw data follows:

Observations of a subject taking a multiple choice test...

> Subject appeared to read passage carefully, then read the question stem and skimmed the passage. Thereafter, eye movements were observed to go back and forth from questions to passage until an answer was marked.

Retrospective protocol of a subject after taking a multiple choice test...

1. well, basically I read the passage
2. and then (I) went back and forth
3. I remember the first question
4. (it) was about what happened in the first paragraph
5. because there was a quote
6. so I looked
7. and it was there
8. I found the right answer
9. I look at all the choices
10. I see if the answer I think is right is there
11. then I mark the right answer

Introspective protocol of a subject taking a multiple choice test...

1. well, I am reading the passage
2. it happened in the 1950's
3. it happened again in the 1970's
4. wait, I need to read this again
5. it's about an iceberg and
6. towing it somewhere
7. I can just see a ship with an iceberg behind it
8. people didn't think they could do it

Levels of Comprehension

Since this research is directly related to the testing of comprehension, it seemed appropriate to score the reading comprehension tests which had been given.

The multiple choice test were scored in the traditional manner. Each correct response was counted as one point. Percentages of correct items were calculated to determine the final score. The average score on the multiple choice test was 92% correct.

As is traditional with strict cloze scoring, only exact replications of the author's words were counted as acceptable in the scoring of the cloze procedure (Taylor, 1953). One point was given for each correct response. Percentages of correct responses were calculated to determine the final score. The average score on the cloze tests was 61%. These scores reflect the fact that only exact word replacement was counted as correct, and is in fact a high score for such a test.

The scoring system used to score the written retellings was modified slightly from one developed by Smith and Jackson (1985). The researcher and two graduate students read each passage independently. Each created a statement of the main idea and listed what they thought were the significant details and supporting details. Significant details were defined as those that

were vital to the understanding of the content. Supporting details were defined as those that supported the significant details. There was 100% agreement between the three raters (researcher and two graduate students) as to the main idea of each passage. There was a 93% agreement on the significant details and 84% agreement on the supporting details. These differences were discussed until agreement was reached for the significant and supporting details.

In the actual scoring of the retellings, five points were given for a statement of the main idea, three points were given for each significant detail included in the retelling, and one point was given for each supporting detail. In addition, a score was given for the overall organization and completeness of the retelling. This score was based on the following scale:

0 = no representation of an overall structure
5 = incomplete representation of an overall structure
10 = reasonably complete representation of an overall structure
15 = strong representation of an overall structure
20 = very strong representation of an overall structure

After a training session, the same two graduate students who had assisted with the main idea and detail identification served as raters for the retellings.

Each passage had a possible total score of 20 points for the overall structure, 5 points for a statement of the main idea, 3 points for each significant detail from the passage, and 1 point for each supporting detail from the passage. The scores were turned into percentages. The average score for the written retellings was 57%. Since there was no prompting as is often provided in oral retellings, these scores can be considered high.

Data Analysis Procedures

There were two major phases of the data analysis procedures. The first was a qualitative analysis. This phase was an attempt to look holistically at the data, and describe what the subjects appeared to be doing in order to understand the text and complete the task. The second phase was a quantitative analysis. This phase examined in raw numbers the different reports of processes by subject and determined statistically significant differences of processes by condition.

Phase I

The Qualitative Analysis

This phase of the analysis was conducted in four steps. The first step was to summarize the researcher's notes from the observation-only interviews. The second and third steps involved the verbal protocols. The second step was to identify and categorize the reading processes used by subjects to complete the tasks. The third step was to determine which processes were employed by each subject across all tasks and which processes were employed to complete each task across all subjects. The fourth step was to summarize the responses given in the exit interviews.

Step 1: Observation-Only Analysis

Each subject was observed by the researcher during the first four interviews. During these initial interviews, subjects were not asked to report what they were doing. The researcher kept notes of her observations. These sessions were not video-taped, therefore validation of her observations was not possible. However, an examination of the researcher's notes by an independent reader revealed that subjects had demonstrated noticeably similar behaviors by task. Summaries were written from the researcher's notes of

the behaviors that had been observed of all subjects under each condition.

Step 2: Emergence of Categories

This step of the analysis was conducted to categorize the reading processes reported by the subjects in the verbal protocols. The researcher carefully read all of the protocols by subject and composed a list of what the subject appeared to be thinking and doing in order to understand the text and complete the tasks. For this analysis, no regard was given to the type of task the subject was engaged in.

Next, the researcher examined the protocols by task. For example, all of the protocols from the multiple choice tests were read as a group. A list was composed of what the subjects as a whole appeared to be doing to understand the text for the multiple choice tests, cloze tests, written retellings, and nonassessed reading task. This was done in order to get an idea if there were any differences in reading processes according to task.

In order to check for the validity of these lists, the protocols were given to four doctoral students in the Language Education Department at Indiana University who assisted in this step of analysis. These readers did

not read the researcher's lists prior to compiling their own list.

Each independent reader was given all of the protocols by subject. They were asked to write a list of what they thought the subjects were doing to understand the text and complete each task.

Next, each independent reader was given a complete set of the protocols by task. One was given all the protocols from cloze tests, another the written retellings, another the multiple choice tests, and another the nonassessed reading activity. Each then wrote a list describing what they thought the subject's were doing to understand the text and complete each task.

Finally, the lists written by the researcher were compared to those written by the independent readers for validity purposes. Both the lists from the individual protocols and those by task were compared. There were several matches between the lists written by the researcher and of the readers that were apparent. Two examples of these were "rereading" and "using prior knowledge". Other matches appeared to be similar, even though different terminology was used to describe what the subject was doing. An example of this was that one independent reader said that the subject was

"formulating questions" while the researcher had called this "predicting". Differences such as these were discussed between the researcher and the readers until a consensus was reached. The lists written by the researcher were then revised to take into account all differences of opinion from the independent readers.

These lists served as a basis for the categorization of the cognitive processes used by the subjects. There were no a priori determinations of categories, therefore, the nature of the analysis was "emergent". This was done by identifying aspects of the lists which seemed to represent reading processes, such as "rereading" and "using prior knowledge". Once a list of categories was made, the researcher met several times with one of the independent readers to explain and define her definition of each category. As a result of these discussions, two processes that had been identified by the researcher were merged into one and three new processes emerged. Even though there were no a priori determinations of categories, the categories which were identified were very similar to those found in the literature (Pritchard, 1987; Wingenbach, 1984; Baker and Brown, 1980; Olshavsky, 1975). The independent reader, as well as the researcher, were to some degree both familiar with much of this literature.

Therefore, this may have influenced the way in which the processes were categorized. A total of twenty-one categories were identified. A definition of each category was written. Examples from the protocols to illustrate the meaning of each category were identified. A compete list of these categories is presented in Chapter 4. An example of one of the categories follows:

Failure to Understand:

> Subject recognizes a loss of concentration or understanding.
>
> Example: I was just off in space somewhere, then I started concentrating and it finally came to me.

Step 3: Process Identification by Subject and by Task

This step served two functions. First, it involved determining what processes were reported by each subject across all tasks. This was done to obtain an overall picture of each subject as a reader and determine what processes were available to him or her. The verbal protocols were reread for each subject. A list of the processes reported was then created for each of the nine subjects. The second function this step served was to determine what processes were used under each condition across all subjects. This was done to obtain an overall picture of the processes employed across all subjects by

task. The protocols were reread by each task. Lists of reported processes were created for each of the four tasks.

Step 4: Exit Interviews

The exit interviews were transcribed. Because the responses of each subject were relatively short, but varied, it was decided to present them in their entirety. These are presented in Chapter Four.

Phase II

The Quantitative Analysis

A quantitative analysis was performed in order to describe numerically differences in process used by subject and by task. This phase also determined statistically significantly different reports of processes under each condition. This phase of the analysis involved two steps.

Step 1: Assignment of Idea Units to Categories

In the first step of Phase II, the researcher assigned each idea unit of the protocols into the categories that had been identified in Step 2 of Phase I. The frequency of the categories were then tabulated. This was done so that the occurrences of

each category could be examined according to subject and to task. Each transcript was read carefully by the researcher. Marginal notes were written on the transcripts assigning each idea unit to one of the categories. Several of these were easily identifiable because the subjects were very explicit as to what they were doing, such as rereading. Approximately eight percent of the idea units were not as easily categorized. After discussing these with two doctoral students in reading education, all but thirty-three idea units were categorized. These were put into a new category named "no specific strategy".

The reliability of the assignments made by the researcher was checked by two elementary school teachers who had taken at least two master level courses in reading education within the past three years. Approximately thirty percent of the protocols were randomly selected. Each rater was given a description of each category and an example from the protocols. After a brief training session, the raters were told to reclassify each response. Their classifications were compared with those of the researcher to determine inter-rater reliability. There was 82% agreement between the researcher's assignments and those of the

raters. The remaining 18% were discussed until agreement was reached.

Step 2: Statistical Tests

In this step of the analysis, cell means and standard deviations were computed for each process category. At this point, it was determined that a nonparametric technique of hypothesis testing (Siegel & Castellan, 1988) was suited to the data for two reasons. First, these tests are often called "distribution-free", meaning that they do not assume that the numbers were drawn from a normally distributed population. Secondly, the tests are considered useful with small samples. Since this study used only nine proficient readers as subjects, this type of analysis was deemed appropriate.

Most nonparametric techniques are "ranking tests." This means that they may be used with scores that are not exact in any numerical sense, but are simply ranks (Siegel & Castellan, 1988). Therefore, a ranking test was determined appropriate to this study. This part of the analysis was based on the occurrence of reported processes. However, the counting of these processes cannot be assumed to be exact, since subjects may have used processes more than what was reported. Therefore, the results of this analysis can only conclude which

processes were reported more than others under each reading task.

A test formulated by Friedman (1937) called the Friedman Two-Way Analysis of Variance by Ranks (Siegel & Castellan, 1988) was determined to be the appropriate statistical test for the data. This analysis can be used to compare the same subjects under different conditions. This analysis tests the null hypotheses that the subjects will perform the same under all conditions. If the null hypothesis is true, the distribution of ranks would be a matter of chance, therefore the ranks would appear with about equal frequency. If the alternative hypotheses is true, then at least one pair of conditions has different average ranks, and therefore the subjects performed differently under at least one pair of conditions. Prior to the computation of this test, a significance level was set at $\leq .05$, a commonly used value of significance (Siegel & Castellan, 1988). This means that if a significant difference is found, there is less than a 5% possibility that the difference was due to chance.

If the obtained value is significant, the Friedman test indicates that at least one of the conditions differs from at least one other condition. A post hoc multiple comparisons analysis between conditions was

conducted to determined which condition or conditions differed from the rest (Siegel & Castellan, 1988).

Limitations

As with any study, this one has several limitations. The limitations are divided into three sections: Sampling Limitations, Instrument Limitations, and Data Collection/Analysis Limitations.

Sampling Limitations

The subjects in this study were probably more familiar with multiple choice tests than with cloze tests or retellings. Therefore, they may have had a more fixed set of strategies for taking multiple choice tests than for cloze tests or giving retellings.

Furthermore, the data represents only proficient sixth grade readers living in a small midwestern city that is the home of a major university. As stated earlier, the small sample size limits the generalizability of the results.

Another limitation involves the fact that the students were aware that the tests they were taking would not count as part of their grade in school. Therefore, they may have taken the tests differently or

not have used the same processes as they would have if the test would have been for credit. However, had the subjects been told that the tests were for credit, they may have been unwilling to supply honest information out of fear that this information might adversely affect their grade.

Finally, because the same subjects were used for each task, there may have been an experimental effect that wouldn't be have occurred had a different subject been used for each assessment task.

Instrument Limitations

Because it was felt that the passages should be as similar as possible to each other, only expository passages were used in this study. Therefore, the results of this study only represent comprehension processes used with expository passages. Further research is needed to determine if there are differences in processes while reading narrative passages.

In addition, the passages were relatively short. Future research should consider using longer passages to determine if passage length is significant.

Data Collection/Analysis Limitations

The most serious limitation of this study may be the data collection procedure. What students say may differ from what they do. Measuring reading processes through concurrent verbal reports may distort what the reader would do under normal conditions. Measuring reading processes through retrospective verbal reports may be invalid because subjects may forget or unable to verbalize their mental processes.

Another limitation is that proficient readers may be too flexible and adaptive to allow researchers to capture their skills in a small sample of situations and options. This loss-of-awareness phenomenon, labeled the "paradox of expertise" by Johnson (1984) and "tacit understanding" by Polanyi (1973), refers to an expert's inability to verbalize a process they engage in without conscious attention. For many readers, these strategies appear to operate at an unconscious, automatic level inaccessible to verbalization or even reflection.

At this time, there is really no way to avoid these rather serious limitations when investigating the reading process. Concurrent and retrospective verbal reporting remain as the most widely-accepted methods of measuring mental processes. Furthermore, it is believed that the inclusive use of observation, as well as both

concurrent and retrospective verbal reports, alleviated some of the problems discussed in this section.

Every effort was made to insure that the setting was similar for each session. The directions that each student received prior to the task were carefully defined and remained consistent for each interview. However, occasionally there were unavoidable interruptions, such as fire drills or teachers/students entering the room unannounced. When these types of unavoidable interruptions occurred, every effort was made to return to the task as quickly as possible.

Although it didn't appear to affect the interviews, the fact that five of the interviews were held in the afternoon instead of in the morning may have affected the readers' verbalizations of their reading processes. There was a loss of information due to replacing the frequency of process with their rank values in the quantitative analysis. As stated earlier, the counting of the occurrences of reported processes is at best an ordinal scale. The researcher cannot assume that these counts were exact or that there was an exact interval between numbers. The subjects may have used processes more often than were reported, or used processes which were unreported. However, it is felt that this loss is offset by the prospect of being able to formulate a

legitimate statistical procedure for testing a hypothesis of interest without having to justify untenable distribution assumptions.

Although validity and inter-rater reliability checks were performed on the verbal protocols and the retellings, it is possible that the researcher's biases or those of the independent readers may have entered the final analysis.

Summary

Despite the reservations associated with this line of investigation, this research provides useful insights into how a testing task affects reading processes. These procedures produced a great deal of information about the reading processes of nine proficient readers and how those processes differed under different testing formats.

CHAPTER IV

RESULTS

This study investigated how reading processes differed as subjects engaged in a multiple choice test, a cloze test, a written retelling, and a nonassessed reading activity. Nine proficient sixth-grade readers participated in the study. During the first four interviews subjects were observed as they completed each task. After the observation-only interviews, verbalizations were collected as the subject engaged in each task (concurrently) and after each subject completed each task (retrospectively). Data analysis consisted of two phases, a qualitative analysis and a quantitative analysis. This chapter presents the results of these two phases as described in Chapter Three.

Phase I

The Qualitative Results

This phase of the analysis was conducted in four steps. The first was to summarize the researcher's notes from the observation-only interviews. The second and third steps involved the verbal protocols. The second step was to identify and categorize the reading processes used by subjects to complete the tasks. The third step was to determine which processes were employed by each subject across all tasks and which processes were employed to complete each task across all subjects. The fourth step was to summarize the exit interviews.

Step 1: Observation-Only Results

During the first four interviews with each subject, the researcher observed the eye and hand movements of the subject as the task was completed. The eye and hand movements of the subjects were very similar when compared by task. The following summaries express the observations that were common among all subjects by task.

Figure 2

Observation-Only Summaries

Multiple Choice:

> Subjects would read through the passage with some back and forth eye movements. When answering questions, subjects often went back to the text.

Cloze Test:

> Subject's eyes would almost always read past each blank before any attempt to fill it in. Sometimes subjects would read several sentences ahead before attempting to fill in a blank. Usually subjects would reread entire sentence after filling in a blank.

Written Retelling:

> Subject would read entire passage slowly. Usually subjects would read passage a second time. Before writing, subject's eyes would linger over certain portions of the text.

Nonassessed Task:

> Subjects read through article quickly, some back and forth eye movements were observed.

Step 2: Categories of Processes

The purpose of this step was to identify and define categories of cognitive processes. The researcher created lists of the cognitive processes each subject appeared to be using. These lists were compiled by subject and by task. The validity of these lists were checked by four independent readers. These readers were asked to create lists of reading processes from the

verbal protocols without having read those written by the researcher.

Once these lists were validated and revised, categories of processes were identified. These categories are presented with a description of that they mean and an example from the protocols. These will be presented in two sections. First are those processes which were only reported in the testing conditions. These are presented under Figure 3 as Test-Taking Processes. A few of these processes have examples from the protocols of different testing conditions. This is because although the idea units seemed to fit together under one category, there may be a slight difference in the way they were used according to task. The second section presents processes which were reported both in the testing conditions and the non-assessed task. These were considered to be more general reading processes, therefore they have been labeled Reading Processes. These are presented under Figure 4.

Figure 3

Test-Related Processes

RECALL

Subject recalls something they read in the passage.

Example: "I remember the two reasons why they were on the rig."

KEY WORD

Subject refers to using certain words to complete task.

Examples:

Multiple Choice:
"I'm looking back at the passage trying to find the word 'breaks' because the question is asking about when icebergs break."

Written Retelling:
"I am trying to remember some of the important words."

ELIMINATE

Subject is able to eliminate possible answers.

Examples:

Multiple Choice:
"I can choose the ones (answers) that look reasonable. A lot of the time I can tell some aren't right. I can usually get it down to two choices."

Cloze Test:
"I can usually think of three or four words that could go in the blank, but then I just choose the one I think sounds best."

READ AHEAD

Subject makes a reference to reading ahead of the blank to determine what word will fit.

Examples:

Cloze Test:
"I'm reading the next few sentences, they may help me decide what will fit here."

SPELLING

Subject shows concern over the correct spelling of a word.

Examples:

Cloze Test:
I'm not sure how to spell "continued", but I'm pretty sure that is the word.

Written Retelling:
I can't remember how to spell "Kieland".

GUESS

Subject decides they will have to guess at an answer.

Example: I guess I will have to take a guess, I can't figure it out.

JUDGE IMPORTANCE

Subject states that he or she believes something is important and will try to remember it.

Examples:

Multiple Choice:

I went back and read the Boston stuff, that was important.

Written Retelling:

I didn't read that part as carefully, I don't think there was anything too important there.

MAIN IDEA

Subject states that he or she is trying to find the main idea.

Example: I'm trying to decide what this is all about so I will know the main idea.

ORGANIZATION

Subject makes a reference to the organization of the text.

Examples:

Multiple Choice:

The first question is usually in the first part of the story, then the next question is usually in the next part of the story and so on. That doesn't always work, but it usually does.

Written Retelling:

I'm trying to see how this is organized, so I can write it down the way it is.

SKIP

Subject decides to leave an answer blank.

Example: I can't figure it out, so I'm going ahead.

CONTEXT CLUES

Subject directly refers to syntax, or states that the word just came to him or her. This process was only reported in the cloze test. It was assumed that when students reported a word "just coming to them" as they read the cloze test, that they were using context clues.

Example: This must be an adjective.

or

I don't know how I got the word, it just popped into my head.

Figure 4

General Reading Processes

REREAD

Subject reads something again.

Examples:

Multiple Choice:
I'm going to look back in the passage, I'm not sure what the answer is.

Cloze Test:
I am reading this sentence over again because I'm not sure what they answer is.

Written Retelling:
I am reading this again, I think this part is important.

Nonassessed Task:
I didn't understand that, I'm going to read it again.

PARAPHRASE

Subject puts text into his or her own words.

Example: They (the turtles) couldn't eat without their long necks.

PRIOR KNOWLEDGE/EXPERIENCE

Subject relates prior knowledge or experiences to the text.

Example: I was thinking about my friend and I and when we play monopoly over at her house.

OPINION

Subject forms an opinion about the text.

Example: I don't think these violets are your everyday delicacies.

PREDICTION

Subject makes a prediction about what will happen or formulates a question.

Examples: I think it (the ship) isn't going to make it. They say it was unsinkable, that may mean trouble.

I'm thinking what does a sailboat have to do with a bike?

CONFIRM/DISCONFIRM

Subject confirms or disconfirms a prediction.

Example: Yes, it did sink.

or

I was wrong, the turtles aren't extinct yet.

RATE

Subject refers to changing the rate of reading.

Example: I'm reading this part more carefully.

or

I just read it pretty fast.

FAILURE TO UNDERSTAND

Subject recognizes a loss of concentration or understanding.

Examples: I was just off in space somewhere, then I started concentrating and it finally came to me.

or

I didn't understand what it was trying to say.

SPECULATES

Subject makes a speculation beyond text.

Example: It must have cost her parents a bundle to have her in all those sports.

VISUALIZES

Subject makes a reference to trying to visualize what is happening in the text.

Example: Whenever they mention dates from a long time ago I try to think of how it would look. I'm thinking about how this Bakker Street must have looked in the 1800's.

NO SPECIFIC STRATEGY

Subject makes a statement that doesn't relate to the text or reading processes.

Examples: I'm just reading it now.

I wonder where the fire is [after hearing a fire siren].

Step 3: Processes by Subject and by Task

This step served two functions. First, it involved determining what processes were reported by each subject across all tasks. Secondly, it determined what processes were reported under each condition across all subjects.

After the categories had been established, the protocols were again examined by the researcher. All protocols were read by subject. A list was created of all the processes used by each subject at least once. These lists follow:

I. During at least one of the tasks, all subjects reported:

a. recalling information from the text
b. eliminating possible answers
c. guessing at answers
d. using context clues
e. rereading portions of the text
f. tieing prior knowledge in with the text
g. adjusting their rate of reading
h. visualizing what was happening in the text

II. Six processes were reported by all but one subject during at least one of the tasks. These include subjects' reports of:

a. using key words to determine correct answers
b. judging the importance of portions of the text
c. skipping questions to come back to later

d. paraphrasing the text
e. expressing opinions of the content
f. predicting what would happen next

III. Three processes were reported by all but two subjects during at least one of the tasks. These included reports of the following:

a. expressing concerns about proper spelling
b. attempting to identify the main idea
c. stating a failure to understand

IV. The remaining four categories of responses were not reported by three or more subjects. They included reports of the following:

a. reading ahead
b. commenting on the organizational structure of the text
c. speculating beyond the text

Table #3 represents the processes which were reported by each subject.

Table #3

Cognitive Processes Reported by Each Subject

	SUBJECT								
PROCESS	1	2	3	4	5	6	7	8	9
Test-Related:									
Recall	X	X	X	X	X	X	X	X	X
Key Word	X	X	X	X	X	X	X	0	X
Eliminate	X	X	X	X	X	X	X	X	X
Read Ahead	X	0	X	0	X	X	X	0	X
Spelling	X	X	0	X	X	X	X	0	X
Guess	X	X	X	X	X	X	X	X	X
Judge Importance	X	0	X	X	X	X	X	X	X
Main Idea	X	X	X	X	0	X	X	X	O
Organization	0	X	X	0	0	0	X	X	X
Skip	X	0	X	X	X	X	X	X	X
Context Clues	X	X	X	X	X	X	X	X	X
General Reading:									
Reread	X	X	X	X	X	X	X	X	X
Paraphrase	X	X	X	X	X	X	X	0	X
Prior knowledge	X	X	X	X	X	X	X	X	X
Opinion	X	X	X	X	0	X	X	X	X
Prediction	X	X	X	X	X	X	X	0	X
Confirm/Disconfirm	X	X	0	X	X	X	0	0	X
Rate	X	X	X	X	X	X	X	X	X
Fail to Understand	X	X	X	X	X	X	0	0	X
Speculates	0	X	X	X	X	0	0	X	0
Visualize	X	X	X	X	X	X	X	X	X
TOTALS	19	18	19	19	18	19	18	14	19

X = Process Reported
0 = Process Unreported

All of the protocols were also examined by the researcher by task. All protocols were read by task. A list was created of all the processes reported under each condition (multiple choice, cloze test, written retelling, and nonassessed task). This examination led to the results presented below.

I. Under all four conditions, at least one subject reported using all of the processes that had earlier been determined to be reading as opposed to test-related processes. These include reports of the following:

   a. rereading portions of the text
   b. paraphrasing the text
   c. tieing prior knowledge in with the text
   d. expressing an opinion of the content of the text
   e. predicting what would happen next
   f. confirming/disconfirming predictions
   g. adjusting their rate of reading
   h. stating a failure to understand the text
   i. speculating beyond the text
   j. visualizing what was happening in the text

II. Two processes were reported across all testing conditions (multiple choice, cloze test, and written retelling). Under these conditions, at least one subject reported the following processes:

   a. recalling information from the text
   b. guessing at answers

III. Two processes were reported only in the multiple choice test and the cloze test. Under these conditions, at least one subject reported:

a. eliminating possible answers
b. skipping questions to come back to later

IV. Two processes were reported only in the cloze test. They were:

a. reading ahead
b. using context clues

V. Four processes were reported only in the multiple choice test and the written retelling. They were:

a. using key words to aid in answering questions
b. judging the importance of portions of the text
c. attempting to identify the main idea
d. commenting on the organization of the text

VI. One process was reported only in the cloze test and the written retelling. It was:

a. expressing a concern over proper spelling

Table #4 represents the processes which were reported by task.

Table 4

Cognitive Processes Reported by Task

	TASK			
PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	X	X	X	0
Key Word	X	0	X	0
Eliminate	X	X	0	0
Read Ahead	0	X	0	0
Spelling	0	X	X	0
Guess	X	X	X	0
Judge Importance	X	0	X	0
Main Idea	X	0	X	0
Organization	X	0	X	0
Skip	X	X	0	0
Context Clues	0	X	0	0
General Reading:				
Reread	X	X	X	X
Paraphrase	X	X	X	X
Prior Knowledge	X	X	X	X
Opinion	X	X	X	X
Prediction	X	X	X	X
Confirm/Disconfirm	X	X	X	X
Rate	X	X	X	X
Fail to Understand	X	X	X	X
Speculate	X	X	X	X
Visualize	X	X	X	X
TOTALS	18	17	17	10

MC = Multiple Choice Test
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading

X = Process Reported
0 = Process Unreported

Step 4: Exit Interviews

The responses to the exit interviews were as follows:

Figure 5

Responses to Exit Interviews

S# = Subject number

Question #1

How did you go about completing the multiple choice test?

S1: On the multiple choice test I knew I could go back to the story to find the answer if I had to. I just sort of read it.

S2: On the multiple choice I read the story carefully. When I got to the answer, I eliminated some of them. I would get it down to two or three answers that sounded the best. Then I went back to the story and found out what was the best answer. If I couldn't find it, I just marked the one that sounded the best.

S3: These tests were easy. When it gives you the answers you know that one of them must be right. I just paid attention as I read the story.

S4: On the multiple choice I just read the passage pretty carefully. Then I read the questions and all the choices. If I remembered the answer I just put it down. If I didn't remember it I would look back.

S5: I just read the passage. Then I would look at the questions and then go back to the passage if I had to.

S6: On the multiple choice test I just read the story pretty carefully. Then I looked at the questions. Usually I would go back and make sure the answer I thought was right was right.

S7: On the multiple choice I tried to read the highlights. That's usually what the questions will ask about. I just try to get the main subject and important details.

S8: I read it pretty carefully, then I answered the questions. I would go back and look at the story if I had to.

S9: On the multiple choice I just read the passage and then went to the questions. Sometimes I knew the answers. If I didn't I would go back and read the passage again.

Question #2

How did you go about completing the cloze test?

S1: I think I had to read the cloze test the most carefully. I kept having to read the whole thing again and again. I couldn't just like write the word "cruise", I would have to keep reading it over and over until it made sense.

S2: On the cloze test I read the passage carefully. I filled in the blank with the word that I thought was the best answer. I just kept reading until I could think of a word that sounded right. I left some of them blank if I couldn't get them until the end.

S3: The cloze test was real hard. I just thought through it. Thought about what word might fit there.

S4: On the cloze test I usually would have a word just come to my head that would be natural. If a word didn't come naturally I would keep reading it and guess at a word.

S5: On the cloze test I would always try a few words. Then I would just choose the one I thought was right.

S6: On the cloze test I just kept reading the sentence. I could usually think of three or four words that might fit in.

S7: On the cloze test, you just use your common sense. If you are writing a story, what word would you put in there?

S8: On the cloze test I just kept reading the sentence until I could get a word that sounded right.

S9: The cloze test was hard. I just had to keep reading them until I would get a word in my head. Some were really hard. I had to keep reading the sentence over and over. Others just seemed to come to me without much trouble.

Question #3

How did you go about completing the written retelling?

S1: The written retelling is the one that I thought I had to get the whole main idea. I would have to memorize it more than I'm used to. I couldn't go back and look at it.

S2: On the retellings, I read that very carefully. I tried to remember as much as I could. Then I just wrote it down on the paper.

S3: These were really hard. You just had to know the story. When you were reading you had to try to memorize it all.

S4: On the written retelling I read it carefully. Then I would look back at important words. I think I read it more carefully than the others.

S5: On the written retelling I read the important parts over and over. It was hard to remember the important things.

S6: On the written retelling I just tried to remember the key words. Then I would write down the key words and start writing them in sentence form.

S7: On the written retelling, you just try to decide what the highlights are and then written them down.

S8: On the retelling, it was really hard to remember. I read it not carefully enough I guess, but I tried. My mind would just go blank when I started to write.

S9: On the written retelling, I paid a lot of attention to the important words. I had to read it pretty carefully I guess.

Question #4

Think about the passage you read when you did not take a test. How did you read it? Which test passage did you read the most like like you read this one?

S1: I think I read the multiple choice test the most like this one. I just read it.

S2: On the one I didn't have to take a test on, I just read it pretty carefully. I don't know why. I guess if it was the only thing I had to read I would read it carefully. I think I read the multiple choice test kind of like that.

S3: If I'm not interested in something I just kind of sluff through it. It depended on how much I was interested in how I read these. If I liked it I read it more carefully than if I didn't like it. I don't think I read any of the others that way.

S4: On the one I didn't take a test on I just read it through normal. I don't know how to describe it. I guess I kind of read the multiple choice test that way. I didn't have to read it all that carefully. I just read it pretty carefully, but not like the others.

S5: I just read it through. I think I read the written retelling the most like I did the one where I didn't take a test.

S6: I don't know, I just read it. I think I read the written retelling the most like the one where I didn't take a test.

S7: I just sort of read it kind of carefully. I think I read the multiple choice test that way.

S8: On the one where I didn't take a test, I read it pretty close, but I didn't read it that carefully. I sort of read the multiple choice test that way.

S9: I just read it kind of carefully. I'm not sure what test I read that way.

Question #5

Which test was the hardest? Why?

S1: I think the cloze test was the hardest. You didn't have choices to choose from. It was hard.

S2: The cloze test was the hardest. Some of the answers were hard to figure out. I had to keep going back and trying to find a word that would work. Some of them just didn't seem right. I just had to keep reading it over and over.

S3: The cloze tests are really hard, but the retellings are even worse. You have to memorize the stuff. I don't like to memorize things. At least in the cloze test there were hints.

S4: I think the written retelling was the hardest. It was even harder when I didn't know much about the story. I remember on the turtle one I didn't know much about turtles. That made it even harder than usual. The maple syrup one wasn't so hard, I know lots of stuff about maple syrup.

S5: The written retelling was the hardest. You might remember the stuff, but forget it by the time you start writing.

S6: The hardest test was the cloze. It was hard to think of a word that made sense.

S7: I think the cloze test was the hardest. Maybe the written retellings. You almost have to come up with something out of the blue on those.

S8: The retelling was the hardest, you couldn't look back.

S9: The cloze test was the hardest because you I didn't have any choices. You just didn't have any choices, you had to think a lot harder.

Question #6

Which test was the easiest? Why?

S1: I think the multiple choice test was the easiest because you knew you could go back if you had to.

S2: I think the multiple choice test was the easiest. Maybe I am just used to them.

S3: The multiple choice test was the easiest. You know the answer has to be there.

S4: The multiple choice test was the easiest. I knew I could look back if I had to. It also gave me choices. That makes it easier.

S5: The multiple choice was the easiest. You knew the answer had to be there somewhere.

S6: The multiple choice test was the easiest. You knew you could look back at the story.

S7: The multiple choice is the easiest. You can always look back, and you know one of the answers is right.

S8: The cloze test was the easiest, you didn't have to go back and look again and again. It was fun because you got to make up an answer, you didn't have to choose one that was already given.

S9: The multiple choice was the easiest, you had choices and you could look back.

Question #7

If you were a teacher and you wanted to really see how well your students could read, which test would you give them? Why?

S1: I guess I would give them the written retelling. You really have to understand it to be able to write it all down.

S2: If I were a teacher I would test reading with the retellings. I think the kids would read it more carefully because I would take it away from them. That would make them read it more carefully.

S3: The multiple choice test is the one that I think is the best test of reading. You don't have to worry so much while you are taking it.

S4: If I were a teacher I would use the written retellings. They would have to read it real hard to learn the facts in it.

S5: I would give a multiple choice test if I were the teacher. The cloze test didn't seem to be testing much, you just had to come up with a word. The written retelling was just to hard. The multiple choice test it the best.

S6: If I were a teacher I would use the multiple choice test. You have to really understand it to answer the questions.

S7: If I were a teacher I would use the multiple choice. You have to read it to get the questions. You can't just come up with something out of the blue.

S8: I would give the retelling if I was a teacher, you really had to understand what they were saying.

S9: If I were the teacher I would use the written retellings. You could really tell how well they read it by how many important things they could remember.

Phase II

The Quantitative Results

Phase II was a quantitative analysis. This phase described numerically the number of times each process was reported under each condition and determined statistically significantly different uses of processes according to task. There were two phases for this part of the analysis.

Step 1: Results of Categorization

The purpose of this step was to determine numerically the differences in the number of reported processes by subject and by process. Each idea unit form the verbal protocols was assigned to one of the categories described in Step 2 of Phase I.

In order to illustrate the differences in processes as clearly as possible, the results are presented by subject and by process.

By Subject

The analysis of the processes by subject revealed several interesting results. For example, Subject #1 paraphrased a great deal during the multiple choice test, the written retelling, and the nonassessed reading task, but did not report paraphrasing during the cloze test. He tied prior knowledge in with the text quite often during the written retellings and the nonassessed task, but less often with the cloze test, and even less often with the multiple choice test. The process of speculating beyond the text was not reported under any of the conditions by Subject #1. A profile of each subject is presented in Tables 5-13. These tables illustrate the occurrences of reported processes individually by each subject across all tasks. Table

#14 lists the total number of processes reported by all subjects across all tasks.

Table #5

SUBJECT #1 (M)

PROCESS	MC	CT	WR	NA
Test Related:				
Recall	6	4	2	0
Key Word	2	0	0	0
Eliminate	4	0	0	0
Read Ahead	0	8	0	0
Spelling	0	4	0	0
Guess	0	4	0	0
Judge Importance	0	0	6	0
Main Idea	0	0	4	0
Organization	0	0	0	0
Skip	0	2	0	0
Context Clues	0	10	0	0
General Reading:				
Reread	6	9	4	2
Paraphrase	16	0	14	14
Prior Knowledge	2	6	13	13
Opinion	0	0	1	2
Prediction	2	1	3	1
Confirm/Disconfirm	2	0	0	2
Rate	0	0	3	0
Fail to Understand	0	1	0	2
Speculates	0	0	0	0
Visualize	3	2	2	3

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #6

SUBJECT #2 (M)

PROCESS	MC	CT	WR	NA
Test Related:				
Recall	8	0	16	0
Key Word	7	0	0	0
Eliminate	8	0	0	0
Read Ahead	0	0	0	0
Spelling	0	2	0	0
Guess	6	12	8	0
Judge Importance	0	0	0	0
Main Idea	2	0	6	0
Organization	0	0	2	0
Skip	0	0	0	0
Context Clues	0	14	0	0
General Reading:				
Reread	5	18	8	2
Paraphrase	2	4	0	6
Prior Knowledge	6	1	16	12
Opinion	0	0	0	4
Prediction	5	0	2	3
Confirm/Disconfirm	1	0	0	2
Rate	4	0	4	7
Fail Understand	5	1	0	3
Speculates	0	0	0	2
Visualize	1	0	1	0

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #7

SUBJECT #3 (F)

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	10	1	4	0
Key Word	6	0	8	0
Eliminate	22	0	0	0
Read Ahead	0	6	0	0
Spelling	0	0	0	0
Guess	5	12	0	0
Judge Importance	0	0	7	0
Main Idea	4	0	4	0
Organization	5	0	0	0
Skip	6	6	0	0
Context Clues	0	18	0	0
General Reading:				
Reread	8	15	8	6
Paraphrase	0	0	6	6
Prior Knowledge	10	4	8	12
Opinion	4	2	0	11
Prediction	0	0	3	3
Confirm/Disconfirm	0	0	0	0
Rate	2	4	4	4
Fail Understand	4	4	0	1
Speculates	4	0	0	0
Visualize	0	0	0	10

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #8

SUBJECT #4 (F)

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	6	4	14	0
Key Word	4	0	2	0
Eliminate	8	3	0	0
Read Ahead	0	0	0	0
Spelling	0	4	0	0
Guess	0	12	0	0
Judge Importance	0	0	8	0
Main Idea	0	0	4	0
Organization	0	0	0	0
Skip	0	2	0	0
Context Clues	0	14	0	0
General Reading:				
Reread	8	4	5	0
Paraphrase	5	0	0	6
Prior Knowledge	5	12	8	10
Opinion	2	0	2	6
Prediction	8	4	0	0
Confirm/Disconfirm	5	6	0	0
Rate	0	0	3	0
Fail Understand	0	2	0	0
Speculates	0	3	0	0
Visualize	6	2	0	10

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #9

SUBJECT #5 (F)

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	8	0	8	0
Key Word	2	0	0	0
Eliminate	2	0	0	0
Read Ahead	0	5	0	0
Spelling	0	6	0	0
Guess	4	7	0	0
Judge Importance	5	0	4	0
Main Idea	0	0	0	0
Organization	0	0	0	0
Skip	2	5	0	0
Context Clues	0	17	0	0
General Reading:				
Reread	12	12	2	2
Paraphrase	4	0	4	5
Prior Knowledge	3	0	2	3
Opinion	0	0	4	4
Prediction	4	0	0	4
Confirm/Disconfirm	0	0	0	1
Rate	0	3	2	2
Fail Understand	0	0	1	0
Speculates	2	0	0	0
Visualize	0	0	0	2

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table # 10

SUBJECT #6 (M)

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	3	3	3	0
Key Word	18	0	6	0
Eliminate	12	5	0	0
Read Ahead	0	9	0	0
Spelling	0	0	9	0
Guess	0	5	0	0
Judge Importance	0	0	3	0
Main Idea	0	0	2	0
Organization	0	0	0	0
Skip	0	3	0	0
Context Clues	0	8	0	0
General Reading:				
Reread	3	7	3	1
Paraphrase	15	0	7	4
Prior Knowledge	11	3	9	9
Opinion	3	0	0	3
Prediction	3	3	6	6
Confirm/Disconfirm	2	1	2	2
Rate	0	3	2	0
Fail Understand	0	0	0	1
Speculates	0	0	0	0
Visualize	3	1	4	2

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #11

SUBJECT #7 (M)

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	18	0	8	0
Key Word	3	0	3	0
Eliminate	19	5	0	0
Read Ahead	0	12	0	0
Spelling	0	0	4	0
Guess	1	6	0	0
Judge Importance	0	0	6	0
Main Idea	0	0	2	0
Organization	2	0	8	0
Skip	3	6	0	0
Context Clues	0	11	0	0
General Reading:				
Reread	6	9	7	2
Paraphrase	3	0	6	12
Prior Knowledge	3	1	3	4
Opinion	5	3	2	4
Prediction	2	0	3	4
Confirm/Disconfirm	0	0	0	0
Rate	2	6	2	0
Fail Understand	0	0	0	0
Speculates	0	0	0	0
Visualize	3	0	2	4

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #12

SUBJECT #8 (F)

PROCESS:	MC	CT	WR	NA
Test-Related:				
Recall	3	0	5	0
Key Word	0	0	0	0
Eliminate	14	8	0	0
Read Ahead	0	0	0	0
Spelling	0	0	0	0
Guess	0	6	0	0
Judge Importance	0	0	8	0
Main Idea	2	0	5	1
Organization	6	0	5	0
Skip	2	6	0	0
Context Clues	0	14	0	0
General Reading:				
Reread	12	18	8	2
Paraphrase	0	0	0	0
Prior Knowledge	18	6	14	21
Opinion	3	0	0	6
Prediction	0	0	0	0
Confirm/Disconfirm	0	0	0	0
Rate	2	3	3	0
Fail Understand	0	0	0	0
Speculates	1	0	3	2
Visualize	2	1	2	3

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #13

SUBJECT #9 (F)

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	5	2	4	0
Key Word	4	0	6	0
Eliminate	4	0	0	0
Read Ahead	0	12	0	0
Spelling	0	5	4	0
Guess	5	8	0	0
Judge Importance	0	0	5	0
Main Idea	0	0	0	0
Organization	0	0	6	0
Skip	3	9	0	0
Context Clues	0	8	0	0
General Reading:				
Reread	9	12	7	3
Paraphrase	6	4	8	16
Prior Knowledge	16	5	9	12
Opinion	3	0	0	4
Prediction	0	0	0	2
Confirm/Disconfirm	0	0	0	1
Rate	3	5	4	2
Fail Understand	8	2	3	1
Speculates	0	0	0	0
Visualize	3	0	2	4

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

Table #14

TOTALS

PROCESS	MC	CT	WR	NA
Test-Related:				
Recall	67	14	64	0
Key Word	46	0	25	0
Eliminate	93	21	0	0
Read Ahead	0	52	0	0
Spelling	0	21	17	0
Guess	11	72	8	0
Judge Importance	5	0	36	0
Main Idea	8	0	29	1
Organization	8	0	21	0
Skip	16	39	0	0
Context Clues	0	114	0	0
General Reading:				
Reread	69	104	52	20
Paraphrase	51	8	45	69
Prior Knowledge	184	38	82	96
Opinion	20	5	9	44
Prediction	24	8	20	23
Confirm/Disconfirm	10	7	2	8
Rate	13	24	27	15
Fail Understand	17	10	4	8
Speculates	7	7	3	4
Visualize	21	6	10	38

MC = Multiple Choice
CT = Cloze Test
WR = Written Retelling
NA = Nonassessed Reading Task

M = Male
F = Female

By Process

By looking at the frequency of reported processes by each process, it can be determined how the process was used differently by all subjects under all conditions. For example, the process "key word" is clearly a test-related process since it was not reported by any of the subjects in the nonassessed task. The process was used by all but one subject during the multiple choice tests, by all but three subjects during the written retellings, but not by any subjects during the cloze test. Tables 15-35 are presented by reading processes.

Table #15

RECALL

Subject	MC	CT	WR	NA
1 (M)	6	4	2	0
2 (M)	8	0	16	0
3 (F)	10	1	4	0
4 (F)	6	4	14	0
5 (F)	8	0	8	0
6 (M)	3	3	3	0
7 (M)	18	0	8	0
8 (F)	3	0	5	0
9 (F)	5	2	4	0

Table #16

KEY WORD

Subject	MC	CT	WR	NA
1 (M)	2	0	0	0
2 (M)	7	0	0	0
3 (F)	8	0	8	0
4 (F)	4	0	2	0
5 (F)	2	0	0	0
6 (M)	18	0	6	0
7 (M)	3	0	3	0
8 (F)	0	0	0	0
9 (F)	4	0	6	0

Table #17

ELIMINATE

Subject	MC	CT	WR	NA
1 (M)	4	0	0	0
2 (M)	8	0	0	0
3 (F)	22	0	0	0
4 (F)	8	3	0	0
5 (F)	2	0	0	0
6 (M)	0	9	0	0
7 (M)	19	5	0	0
8 (F)	14	8	0	0
9 (F)	4	0	0	0

Table #18

READ AHEAD

Subject	MC	CT	WR	NA
1 (M)	0	8	0	0
2 (M)	0	0	0	0
3 (F)	0	6	0	0
4 (F)	0	0	0	0
5 (F)	0	5	0	0
6 (M)	0	9	0	0
7 (M)	0	12	0	0
8 (F)	0	0	0	0
9 (F)	0	12	0	0

Table #19

SPELLING

Subject	MC	CT	WR	NA
1 (M)	0	4	0	0
2 (M)	0	2	0	0
3 (F)	0	0	0	0
4 (F)	0	4	0	0
5 (F)	0	6	0	0
6 (M)	0	0	9	0
7 (M)	0	0	4	0
8 (F)	0	0	0	0
9 (F)	0	5	4	0

Table #20

GUESSING

Subject	MC	CT	WR	NA
1 (M)	0	4	0	0
2 (M)	6	12	8	0
3 (F)	5	12	0	0
4 (F)	0	4	0	0
5 (F)	4	7	0	0
6 (M)	0	5	0	0
7 (M)	1	6	0	0
8 (F)	0	6	0	0
9 (F)	5	8	0	0

Table #21

JUDGE IMPORTANCE

Subject	MC	CT	WR	NA
1 (M)	0	0	6	0
2 (M)	0	0	0	0
3 (F)	0	0	7	0
4 (F)	0	12	0	0
5 (F)	5	0	4	0
6 (M)	0	0	3	0
7 (M)	0	0	6	0
8 (F)	0	0	8	0
9 (F)	0	0	5	0

Table #22

MAIN IDEA

Subject	MC	CT	WR	NA
1 (M)	0	0	4	0
2 (M)	2	0	6	0
3 (F)	4	0	4	0
4 (F)	0	0	4	0
5 (F)	0	0	0	0
6 (M)	0	0	2	0
7 (M)	0	0	2	0
8 (F)	2	0	5	1
9 (F)	0	0	0	0

Table #23

ORGANIZATION

Subject	MC	CT	WR	NA
1 (M)	0	0	0	0
2 (M)	0	0	2	0
3 (F)	5	0	0	0
4 (F)	0	0	0	0
5 (F)	0	0	0	0
6 (M)	0	0	0	0
7 (M)	2	0	8	0
8 (F)	6	0	5	0
9 (F)	0	0	6	0

Table #24

SKIP

Subject	MC	CT	WR	NA
1 (M)	0	2	0	0
2 (M)	0	0	0	0
3 (F)	6	6	0	0
4 (F)	0	2	0	0
5 (F)	2	5	0	0
6 (M)	0	3	0	0
7 (M)	3	6	0	0
8 (F)	2	6	0	0
9 (F)	3	9	0	0

Table #25

CONTEXT CLUES

Subject	MC	CT	WR	NA
1 (M)	0	10	0	0
2 (M)	0	14	0	0
3 (F)	0	18	0	0
4 (F)	0	14	0	0
5 (F)	0	17	0	0
6 (M)	0	8	0	0
7 (M)	0	11	0	0
8 (F)	0	14	0	0
9 (F)	0	8	0	0

Table #26

REREADS

Subject	MC	CT	WR	NA
1 (M)	6	9	4	2
2 (M)	5	18	4	2
3 (F)	8	15	8	6
4 (F)	8	4	5	0
5 (F)	12	12	2	2
6 (M)	3	7	3	1
7 (M)	6	9	7	2
8 (F)	12	18	8	2
9 (F)	9	12	7	3

Table #27

PARAPHRASES

Subject	MC	CT	WR	NA
1 (M)	16	0	14	14
2 (M)	2	4	0	6
3 (F)	0	0	6	6
4 (F)	5	0	0	6
5 (F)	4	0	4	5
6 (M)	15	0	7	4
7 (M)	3	0	6	12
8 (F)	0	0	0	0
9 (F)	6	4	8	16

Table #28

PRIOR KNOWLEDGE

Subject	MC	CT	WR	NA
1 (M)	2	6	13	13
2 (M)	6	1	16	12
3 (F)	10	14	8	12
4 (F)	5	12	8	10
5 (F)	3	0	2	3
6 (M)	11	3	9	9
7 (M)	3	1	3	4
8 (F)	18	6	14	21
9 (F)	16	5	9	12

Table #29

OPINION

Subject	MC	CT	WR	NA
1 (M)	0	0	1	2
2 (M)	0	0	0	4
3 (F)	4	2	0	11
4 (F)	2	0	2	6
5 (F)	0	0	4	4
6 (M)	3	0	0	3
7 (M)	5	3	2	4
8 (F)	3	0	0	6
9 (F)	3	0	0	4

Table #30

PREDICTING - FORMULATING QUESTIONS

Subject	MC	CT	WR	NA
1 (M)	2	1	3	1
2 (M)	5	0	2	3
3 (F)	0	0	2	3
4 (F)	8	4	0	0
5 (F)	4	0	0	4
6 (M)	3	3	6	6
7 (M)	2	0	3	4
8 (F)	0	0	0	0
9 (F)	0	0	0	2

Table #31

CONFIRM/DISCONFIRM

Subject	MC	CT	WR	NA
1 (M)	2	0	0	2
2 (M)	1	0	0	2
3 (F)	0	0	0	0
4 (F)	5	6	0	0
5 (F)	0	0	0	1
6 (M)	2	1	2	2
7 (M)	0	0	0	0
8 (F)	0	0	0	0
9 (F)	0	0	0	1

Table #32

VARIES RATE

Subject	MC	CT	WR	NA
1 (M)	0	0	3	0
2 (M)	4	0	4	7
3 (F)	2	4	4	4
4 (F)	0	0	3	0
5 (F)	0	3	2	2
6 (M)	0	2	2	0
7 (M)	2	6	2	0
8 (F)	2	3	3	0
9 (F)	3	5	4	2

Table #33

FAILURE TO UNDERSTAND

Subject	MC	CT	WR	NA
1 (M)	0	1	0	2
2 (M)	5	1	0	3
3 (F)	4	4	0	1
4 (F)	0	2	0	0
5 (F)	0	0	1	0
6 (M)	0	0	0	1
7 (M)	0	0	0	0
8 (F)	0	0	0	0
9 (F)	8	2	3	1

Table #34

SPECULATES

Subject	MC	CT	WR	NA
1 (M)	0	0	0	0
2 (M)	0	0	0	2
3 (F)	4	0	0	0
4 (F)	0	3	0	0
5 (F)	2	0	0	0
6 (M)	0	0	0	0
7 (M)	0	0	0	0
8 (F)	1	0	3	2
9 (F)	0	0	0	0

Table #35

VISUALIZES

Subject	MC	CT	WR	NA
1 (M)	3	2	2	3
2 (M)	1	0	1	0
3 (F)	0	0	0	10
4 (F)	6	2	0	10
5 (F)	0	0	0	2
6 (M)	3	1	4	2
7 (M)	3	0	2	4
8 (F)	2	1	2	3
9 (F)	3	0	2	4

Step 2: Results of Statistical Tests

Step 2 consisted of a statistical analysis to determine if there were statistically significant differences in processes by task. The Friedman Two-Way Analysis of Variance by Ranks (Siegel & Castellan, 1988; Friedman, 1937) was used to determine these differences (see appendix). A significance level of $\leq .05$ was established before the analysis. Out of the twenty-one reading processes that had been identified, eleven were found to be statistically different. Seven were processes which had earlier been determined to be test-related because they had not surfaced in the nonassessed reading task. Those were:

Recall	Main Idea
Skips	Eliminate
Context Clues	Guessing
Key Word	

The remaining four processes that contained statistically significant differences were those that had earlier been determined to be more related to general reading processes. Those were:

Reread	Opinion
Paraphrase	Visualize

From this information, a post-hoc analysis was conducted to determine where the differences lay across conditions

(Siegel & Castellan, 1988). These differences are outlined in Figure 6.

It should be noted that the Friedman Test failed to recognize one significant difference. Subjects reported "reading ahead" only in the cloze test. Because the Friedman Test uses ranks in place of raw numbers, the multiple choice, written retelling, and nonassessed reading task all tied in second place in rank. Because it was obvious from the raw data that the difference was significant, the post-hoc analysis was performed on the process "read ahead" and a significant difference was found.

Figure 6

Results of Post-Hoc Analysis

TEST-RELATED PROCESSES

Recall

There were two statistically significant differences. The students reported recalling more in both the multiple choice test and the written retelling than in the nonassessed reading task.

Key Word

Subjects reported using key words more in the multiple choice test and the written retelling than in the cloze test or the nonassessed reading task.

Eliminate

There were two statistically significant differences. Student eliminated answers more in the multiple choice test than in either the written retelling or nonassessed reading task.

Guessing

The subjects guessed at answers statistically significantly more in the cloze test than in the written retelling.

Main Idea

There was a statistically significant difference in the reports of identifying the main idea between the cloze test and the written retelling, and between the written retelling and nonassessed reading task. In both conditions the students reported the process more often in the written retelling than they did in either the cloze test or the nonassessed reading task.

Skips

There was a statistically significant difference between the use of this process between the cloze test and the written retelling, and between the cloze test and the nonassessed reading task. In both comparisons, the students used this strategy more in the cloze test situation than in either written retelling or in the nonassessed reading task.

Context Clues

The students used context clues statistically more in the cloze test than in either the multiple choice or the written retelling.

GENERAL READING PROCESSES

Reread

The subjects reported rereading statistically more often in both the multiple choice and the cloze test than in the nonassessed task.

Paraphrase

The only significant difference was between the nonassessed reading task and the cloze test. The students paraphrased more in the nonassessed reading task than in the cloze test.

Opinion

There were two significantly different differences in expressions of opinion of the text. These were between the cloze test and the nonassessed task and between the written retelling and the nonassessed task. In both cases the subjects reported opinions more in the nonassessed task than in the other two.

Visualize

The students reported visualizing the text much more in the nonassessed reading task than in the cloze test.

Summary

This chapter has presented the results of the data analysis for this study. Phase I was a qualitative analysis which consisted of three steps. Step 1 produced summaries of the observation-only interviews. Step 2 involved the identification of twenty-one process categories. Step 3 determined what processes were reported by each subject across all tasks and what processes were used under each condition across all subjects.

Phase II was a quantitative analysis which consisted of two steps. Step 1 involved assigning each idea unit of the verbal protocols to one of the

categories of cognitive processes. Step 2 involved a test to determine statistically significant differences of process use under each condition and a post-hoc analysis to determine where the differences lay. These findings and their implications for future research will be considered in Chapter 5.

CHAPTER V

CONCLUSION

This study was designed to describe and compare the reading comprehension processes of nine proficient sixth grade readers as they engaged in a multiple choice test, a cloze test, a written retelling, and a non-assessed reading task. Each subject was interviewed on twelve different occasions. Each task was presented to the subject three times. During the first four interviews, the subject's eye and hand movements were observed. On the subsequent interviews, the subjects were instructed to verbalize their introspective thoughts both as they were completing each task (concurrently) and after they had completed each task (retrospectively). The resulting data were then analyzed in two phases. This chapter presents a discussion of each phase of analysis and the conclusions which can be drawn. A final section provides implications for future research.

Phase I

Conclusions from the Qualitative Data

The results of the qualitative analysis were summaries of the observation-only interviews and categories of reading processes which emerged from the verbal reports.

Observation-Only Interviews

The observation-only interviews resulted in summaries of observations that were extremely similar when compared across tasks. This suggests that there were at least superficial differences in how the subjects engaged in completing each task. In addition, these observations added further evidence to the validity of the verbal reports. In the multiple choice testing, all subjects were observed reading the passage first, then reading the questions, and often going back to the text. This overall practice was confirmed in the verbal reports. In the cloze test, subject's eyes were observed going back and forth over the same portion of text several times. This was confirmed in the verbal reports, as subjects reported the process of rereading again and again. In the observations of the written retellings, subjects appeared to read the passage

slowly, then go back over certain parts of the text. This also was confirmed by the verbal reports. Subjects reported reading the text "carefully", then rereading certain parts to get a better idea of what the text said, or perhaps to memorize certain portions. In the nonassessed task, the subjects were observed reading the text very quickly, with some back and forth eye movements over the text. This also was confirmed by the verbal reports. Subjects reported reading "quickly", and only rereading when they were interested in understanding the text better.

Two things can be concluded from the observation-only interviews. First, there were at least some superficial differences in the way subjects engaged in each task. In the multiple choice, written retelling, and nonassessed tasks the subjects read through each passage in its entirety then reread certain parts as the subjects deemed necessary. In the cloze test, subjects reread each sentence several times before reading on.

Secondly, any interference in reading processes that verbal reporting may have caused did not affect, at least on this superficial level, the way subjects undertook the task. That is, subjects verbally reported doing what they had been observed doing.

Categories of Processes

The reading processes which emerged from the verbal reports of the data were put into twenty-one categories. This categorization of processes resulted in two different types. First were those processes which were determined to be general reading processes. These were determined to be so because they submerged in all four conditions (multiple choice, cloze test, written retelling and nonassessed task). The second type were deemed to be test-related because they did not submerge in the nonassessed reading task (See Table #4). Eleven processes were determined to be test related since they did not surface in the nonassessed reading task. The remaining ten were determined to be more related to general reading processes.

It can be concluded that all four tasks involved general reading processes. However, the number of times that these processes were reported under each condition was often large. These differences are examined in Phase II, the quantitative analysis.

The categories of processes identified in this study represent a potentially comprehensive framework of those that were used by the subjects in this study. However, because these categories were developed from the subject's protocols, they cannot be presumed to be

representative of all reading or test-related processes. Although the subjects in this research were only sixth graders, they were considered proficient readers by their teachers. Therefore, the following remarks by Brown (1982) should be taken into consideration:

> ...although mature readers typically engage in comprehension monitoring, it is not usually a conscious experience. When comprehension is proceeding smoothly, good readers proceed as if on automatic pilot until a problem is detected. When some triggering event alerts them to a comprehension failure... the understanding process slows down and becomes planned, demanding conscious effort.

In light of this, there may have been processes which were too automatic to be reported by the subjects, and therefore not included in the categories. Therefore, the failure of some subjects to report some processes cannot be assumed to mean that those processes were not employed at all by those subjects. Another explanation for some subjects not reporting certain processes would be because those processes do not exist in their pools of resources of reading comprehension. These are some of the questions which must go unanswered in this investigation. However, it can be concluded that those subjects who reported a process at least once in the verbal reports had the option of employing that processes under all the conditions (See Table #3).

In spite of the limitations of the verbal report data, the results of this research indicate that these subjects were aware of the measures they took to comprehend a text and were able to verbally report at least part of their reading comprehension processes. Therefore, it can be concluded that verbal reports yield reasonably valid data regarding the processes readers employ to facilitate their understanding of a text.

Exit Interviews

The Exit Interviews were given to each subject after all of the other data had been collected. The responses to the exit interviews (Figure 5) reveals several interesting things. On the first question which asked subjects how they completed the multiple choice test, the responses indicated that the subjects felt competent taking these tests. The subjects were aware that there was one correct answer to each question and that if they were unable to recall the correct response they could reread portions of the passage to help them. These responses are not surprising since it can be presumed that the subjects were very experienced with these types of tests.

On the second question which asked subjects how they completed the cloze test, all of the subjects

recalled rereading each sentence until they could think of a word that would fit. Two subjects described the cloze test as being hard and having to read very carefully. Most of the subjects seemed to recognize that often there was more than one word that could fit into the sentence and that they had to rely on their best judgment in choosing one.

On the third question regarding the written retelling, all of the subjects reported attempts to identify key words or important information that they would then try to memorize or remember. Only one subject reported attempting to identify the main idea.

On the fourth question about the nonassessed reading task, subjects had a hard time describing how they read the passage other than just to say they "just read it" or "just read it normally". When asked what test passage they read the most like the nonassessed passage, five said the multiple choice test, two said the written retelling, one said he wasn't sure, and one said he didn't think he read any of them like the nonassessed passage. None of the subjects reported reading the nonassessed passage in the same way as the cloze test, suggesting that these subjects did view reading and completing the cloze passage as more different than the others.

The fifth question asked subjects which test was the hardest. Six subjects reported the cloze test as being the hardest, three reported the written retelling as being the hardest. No subject reported the multiple choice test as being the hardest.

The sixth question asked which test was the easiest. Eight of the subjects reported that the multiple choice test was the easiest. This could be due to the fact that it is very likely that the subjects were more experienced in taking multiple choice tests than in taking the other forms of tests. One subject said that the cloze test was the easiest. This subject liked the fact that she didn't have to return to the passage again and again in the cloze test in order to determine the answer and that there could be more than one correct answer.

The seventh question asked subjects what test they would use if they were a teacher and really wanted to determine how well their subjects had read a text. Five of the subjects replied that they would use the written retellings. Reasons expressed for this were that the subjects thought students had to read it more carefully, identify what was important, and overall understand the passage better because the text was taken away from them while they were writing. Four replied that they would

use the multiple choice test. Reasons expressed for this response were that the subjects thought the students wouldn't have to worry so much while they were reading because they could look back at the passage for help. Other subjects stated that they would choose the multiple choice test because students would have to understand the passage in order to answer the questions. None of the subjects replied that they would use the cloze test as a way to measure reading comprehension.

Several things can be concluded from the exit interviews. First, it can be concluded that subjects viewed the cloze test as being different from the other tasks. Subjects reported reading the test differently in a general way, that being that they reread each sentence several times. Six of the subjects felt the cloze test was the hardest. However, one thought it was the easiest. None of the subjects reported that a cloze test would be their choice as a teacher to measure reading comprehension.

The responses to the exit interviews lead to the conclusion that the written retellings involved memorizing important portions of the text. Two subjects reported that they thought reading the written retelling passage was most like reading the nonassessed passage. Three subjects felt that the written retelling was the

hardest test. None thought that it was the easiest. Five subjects chose the written retelling as the means of assessing reading comprehension that they would use if they were a teacher. Therefore, it can be concluded that although subjects had somewhat mixed feelings about the written retelling because it involved memorization, at least some of the subjects viewed reading the written retelling passages as being similar to reading the nonassessed passages and some believed it was a good measure of reading comprehension.

The responses concerning the multiple choice test indicate that subjects were comfortable with these tests, possibly because they were more experienced with them. None of the subjects thought this test was the hardest test, eight thought it was the easiest. Five subjects thought they read the multiple choice test the most like they read the nonassessed passage. Four subjects reported that they would use a multiple choice test to assess reading comprehension if they were a teacher. Overall, the responses lead to the conclusion that several of the subjects thought that reading multiple choice tests were not very different from reading the nonassessed passages. In addition, some of the subjects felt that multiple choice tests were good ways to measure reading comprehension.

Phase II

Conclusions from the Quantitative Data

The results of the quantitative analysis were frequencies of reported processes by each subject under each condition. These will be discussed by subject, by task, and by processes.

By Subject

Tables 5-13 presented in Chapter 4 show the frequency of reported processes by subject under each condition. These tables reveal that all subjects were aware of and employed most of the general reading processes during at least one of the tasks. The exceptions to this were that four subjects did not ever report speculating beyond the passage, one subject did not report the confirmation or disconfirmation of a prediction, and one subject never reported a failure to understand.

In the test-related processes, four subjects never made reference to the organization of the text, three did not report reading ahead, three did not report an effort to identify the main idea, two did not express a concern over spelling, one never made a reference to judging the importance of information presented in the text, one did not make a reference to the use or

identification of a key word, and one did not mention skipping a problem question (or blank) until later.

Therefore, it can be concluded that all subjects in this study had at their disposal at least eleven of the twenty-one processes that were identified. Those processes were: rereading portions of the text, paraphrasing, tieing prior knowledge in with the text, formulating opinions about the text, predicting what would happen next, adjusting the rate of reading, and visualizing what was happening in the text.

The following test-related processes were reported by all of the subjects: recalling previously read information, eliminating possible answers, guessing at correct answers, and using context clues.

By Task

The following discussions and conclusions for each task is based on the number of times each process was reported by task (Table 14). This section will examine those numbers that the researcher feels are significant to the understanding of the processes involved in each task.

Multiple Choice: All of the reading processes were reported in the multiple choice test. Tieing prior knowledge in with the text was the most often reported

process, being reported 184 times. Paraphrasing the text was reported 51 times. Rereading portions of the text was reported 69 times.

In the test-related processes, elimination was reported 93 times, recall 67 times. Subjects reported using key words to find the correct answer 46 times. Guessing at the correct answer was only reported 11 times.

The significant reports of tieing prior knowledge in with the text and paraphrasing the text during the multiple choice test signifies that the subjects were cognitively involved in comprehending the text. It also can be concluded that recalling information from the text and eliminating possible answers are significant in answering multiple choice test questions. The task of answering multiple choice questions did not appear to significantly alter the number of reported processes as compared to those reported in the nonassessed reading tasks.

Cloze Test: Of the general reading processes, rereading portions of the text was reported 104 times. Tieing prior knowledge to the text was reported 38 times and paraphrasing the text was reported only 8 times. In the test-related processes, the use of context clues was

reported 114 times, guessing 72 times, and reading ahead 52 times.

The numbers of reported processes in the cloze test suggest that the subjects were not deeply involved cognitively with comprehending the text as a whole. The small number of reports of paraphrasing, visualizing, expressing opinions, and using prior knowledge, and the large reports of rereading each sentence suggests that the subjects were breaking the text into small parts in order to fill in the blanks at the expense of getting more cognitively involved with the text as a whole.

The most often reported test-related process was the use of context clues, which were not reported at all during the other three tasks. This does not suggest to the researcher that context clues were not used in the other tasks. However, the large reporting of context clues in the cloze tests suggests that it is that process which cloze tests emphasize. This is not a surprise as that is often what cloze tests claim to measure.

However, this finding suggests that the strong use of context clues may be at the expense of thinking processes that may be more associated with understanding the overall meaning of the text. An important aspect in discussing the cloze test is examining those processes

which were not reported at all, such as judging the importance of portions of the text, identifying the main idea, or referring to the organization of the text. The lack of any report of these processes again suggests that subjects did not appear to be concerned with understanding the text as a whole.

Written Retelling: In the written retelling, subjects reported tieing their prior knowledge in with the text 82 times. Rereading was reported 52 times, and paraphrasing 45 times.

In the test-related processes, not surprisingly, recall was reported a total of 64 times. Judging the importance of information in the text was reported 36 times. Attempting to identify the main idea was reported 29 times.

The reports of paraphrasing, identifying the main idea, and using prior knowledge suggest that the subjects were cognitively involved with the text. The large number of reports of judging the importance of portions of the text suggests that the subjects paid a great deal more attention to text structure in the written retelling than in the other tasks. Subjects possibly saw this as a way to recall information from the text when they were asked to write down what they remembered.

Nonassessed Reading Task: In the nonassessed reading tasks, tieing the text in with prior knowledge was reported 96 times and paraphrasing 69 times. Visualizing what was happening in the text was reported 38 times.

The large reports of the use of prior knowledge, paraphrasing and visualizing under this condition suggests that subjects were involved cognitively with the text.

The absence of a testing condition did leave all of the processes that had been classified as test-related unreported. A few of these are worth noting. Subjects did not report a concern with the main idea or the organization of the text. Therefore, it is interesting to note that these two processes, which are often emphasized in instruction, were not reported at all in the nonassessed reading task.

By Process

Tables 15-35 present the number of times each process was reported by each subject according to task. In the statistical analysis, ten processes were identified as having statistically significant differences by task. Five of those statistical differences were senseless from a practical standpoint.

For example, a significant difference was found under the process recall between the multiple choice test and the nonassessed task. The reason for this difference is obvious since subjects were not asked to recall the text during the nonassessed task. The differences that were deemed to be senseless from a practical standpoint were in the processes of recalling, eliminating answers, guessing, skipping, and using context clues. These differences are included in Figure 3 for examination. This discussion will only cover those processes which were determined to be statistically significantly different in addition to being significant from a practical view.

A note of caution is worthy of mention at this point. The small sample size, and thus the small numbers of reported processes, in the data makes any statistically significant differences suspect. That is why only raw numbers were presented in Chapter 4. In spite of the senseless statistically significant differences mentioned above and the small numbers on which all of the statistical data resulted, the statistical analysis of the data did provide some interesting information. These are presented below with the conclusions that the researcher feels can be made from them.

Key Word

Subjects identified a key word statistically significantly more in the multiple choice test and the written retelling than in the cloze test or the nonassessed reading task.

The fact that subjects reported identifying key words more in the multiple choice test and the written retelling is sensible although the use of the key word that was identified was used differently under the two conditions. In the multiple choice test, subjects identified a key word from the stem to help them locate the answer in the passage. In the written retelling, subjects tried to memorize the key words so that they could reproduce them in their retellings.

Main Idea

There was a statistically significant difference in the reports of identifying the main idea between the cloze test and the written retelling, and between the written retelling and nonassessed reading task. In both conditions the subjects reported the process a great deal more in written retelling than they did in either the cloze test or the nonassessed reading task.

The reporting of finding the main idea more in the written retelling than in the cloze test and nonassessed reading task is logical. Subjects knew that they were going to be asked to reproduce the text in the written retelling and therefore attended to finding the main more than in the other two tasks.

Reread

The subjects reported rereading statistically more in both the multiple choice test and the cloze test than they did in the nonassessed task.

As mentioned in the examination of the cloze test, subjects reread a great deal while completing this task. Almost every sentence was reread at least once. The rereadings were almost always stated attempts to determine what word should go into the blank or checking to make sure that what they determined would go in the blank made sense.

Rereading in the other two tasks was reported much less frequently. Rereading in the multiple choice test was usually expressed when subjects went back to the passage to find or check an answer. Rereading in the nonassessed task was always stated as an attempt to understand more clearly what was being read.

Paraphrasing

The only significant difference was between the nonassessed reading task and the cloze test. The students paraphrased more in the nonassessed reading task than in the cloze test.

Paraphrasing was reported much less often in the cloze test than in all of the other three tasks. This fact signifies that the subjects were not attending to putting the text into their own words because the task of filling in the blanks demanded too much of their

attention.

The fact that students paraphrased most often in the nonassessed task has two possible explanations. First, the subjects may have paraphrased more because of the lack of anything else to report to the researcher. The alternative explanation is that subjects were more interested in or felt they had more freedom to take the time to put the text into their own words. Or, it is also possible that paraphrasing may be a significant process leading to a real understanding of the text, which may have been what the subjects viewed as their goal in the nonassessed task.

Opinion

There were two significantly different differences in expressions of opinion of the texts. These were between the cloze test and the nonassessed task and between the written retelling and the nonassessed task. In both cases, the subjects reported opinions more in the nonassessed task than in the other two.

The fact that subjects expressed opinions more in the nonassessed task statistically significantly more often than in the cloze test and the written retelling again has three possible explanations. As was mentioned above under paraphrasing, perhaps the subjects expressed opinions more in the nonassessed task for the lack of anything else to report to the researcher. The

alternative to that is again that perhaps subjects felt freer to attend to formulating opinions because they were not being assessed. Or it is possible that formulating opinions may be part of trying to really understand a text when it is read for the what the subject may perceive as his or her own purpose for reading.

Visualize

> The students reported visualizing the text much more in the nonassessed reading task than in the cloze test.

Once more, a large difference exists between the cloze test and the nonassessed reading task. Again, subjects may have reported visualizing for the lack of something to say to the researcher, or visualizing the text is used more often when a reader has his or her own purpose for reading as may have been the case in the nonassessed reading task.

Read Ahead

> As reported in Chapter 4, the Friedman Test failed to recognize reading ahead as statistically significant because of the rankings. Reading ahead was only reported in the cloze test.

This difference is logical due to the definition of reading ahead as used in this study. Subjects would leave blanks empty in the cloze test

and read ahead hoping to gather more information before they attempted to fill in the blanks.

Overall Conclusions

The subjects of this study, nine proficient sixth grade readers, expressed flexibility in reading processes and the knowledge of several reading processes.

It can also be concluded that all four tasks; multiple choice tests, cloze tests, written retellings, and nonassessed task, involved processes that have long been considered as associated with reading comprehension.

However, the overall conclusions of this investigation also indicate that the reading processes of the subjects in this study did differ while engaged in each task. The task which stood out as the most different from the other three was the cloze test (See Table #14). The subjects in this study did not appear to be as involved cognitively with the text in the cloze test as they did in the other three tasks. This does not lead the researcher to conclude that the cloze test did not test reading, but it can be said that the task certainly altered reported reading processes.

Therefore, cloze tests may lack construct validity when used to measure general reading comprehension abilities.

These tests seem to be good measures of student's abilities to use context clues, which is one factor in reading comprehension. However, in the opinion of the researcher, they do not "measure what might be called minimally inferential comprehension" as the DRP Handbook (1986) claims. Ashby-Davis' (1985) assertions that cloze reading is not like ordinary reading are supported by the results of this investigation. The results of this investigation support her claims that reading speed, eye movements, and overall reading strategies do change during cloze testing.

The multiple choice test and the written retellings, on the other hand, were very similar numerically to each other and to the nonassessed reading task. In the multiple choice test and the written retelling, the task itself did not appear to upset the reported reading processes to any great degree. Therefore, these tests do appear to involve much the same processes as the nonassessed task. Those processes that were reported most often in taking the multiple choice test, the written retelling, and the nonassessed reading task are those that are considered to be associated with the construct of reading, such as tieing

prior knowledge in with the text, visualizing what is happening in the text, and paraphrasing the text.

When given in the traditional manner, multiple choice tests and written retellings do not provide information about the processes students use to complete them, only the product of those processes. Taken in this light, one could agree with Valencia and Pearson (1987) that "The tests used to measure reading achievement do not reflect recent advances in our understanding of the reading process". However, this investigation provides evidence that these tasks in and of themselves do not significantly alter the cognitive processes generally associated with reading. Therefore, within the limitations of this investigation, it appears that both multiple choice tests and written retellings are possessive of construct validity.

Written retellings do appear to "give us a sense of <u>how</u> as well as <u>how much</u> information is represented in the student's thinking right after reading as Smith and Jackson (1985) claim. The results of this investigation also compliment the statements made by Kalmbach (1986) that retellings reveal:

> (1) The point or points students see in the stories they read; and (2) The problems students have organizing the different elements of a story into a coherent whole (327).

However it is also true that written retellings measure the student's ability to write in addition to their ability to read. Therefore, it is possible that oral retellings may be more valid measures of reading comprehension than written retellings.

The findings surrounding the processes reported while taking the multiple choice test support those of Wingenbach (1984). Wingenbach's categories of strategy use reported during a multiple choice test are very similar to the categories of processes which were identified in this investigation. The findings of this investigation do not however agree with those found by Alvermann and Ratekin (1982). These researchers claimed that subjects in their study reread more in preparation for an essay test than a multiple choice test. This investigation found the opposite to be true, that subjects reread more in the multiple choice test than in the written retelling. There may be a perceived difference among subjects in both studies between preparing for an essay test as compared to preparing for a written retelling. Alvermann and Ratekin also claimed that the seventh and eighth grade average readers in their study may only have a limited awareness of the entire range of strategic activities available. This investigation provides evidence to the contrary, that

sixth grade proficient readers had a rather large pool of processes available to them. These differences in results may be due to the fact that the populations were of different ages and at different levels of proficiency. It should also be noted that Alvermann and Ratekin only employed retrospective interviews in their study. The subjects in their study may have forgotten the processes they employed in completing each task. Further research is needed to determine if this is true.

In addition to the above conclusions, this investigation adds further evidence to the validity of verbal reporting of reading processes. When collected with care and analyzed with caution, they provide a useful way of examining reading processes.

Implications

Even considering the limitations of this research as outlined in Chapter Three, this study has provided insights into how different assessment tasks alter reading processes. The findings provide implications for future research and pedagogy.

Research

Studies with larger populations are certainly needed in this area. Unfortunately, investigations such

as these are very time consuming, and thus difficult to undertake with large populations. Until such studies are conducted, there is really no way that the findings of this research or any other such small investigation can be generalized to a larger population.

Test developers or other interested researchers need to conduct research that investigates whether or not the type of questions used in multiple choice tests alter reading processes. The findings of this investigation did not reveal information on this question. It is possible that in the rereadings of the text, subjects read differently when looking to find the main idea than when they looked to find a detail or make an inference.

A review of similar studies shows that processes tend to differ by age (Garner & Reis, 1981), proficiency (Garner & Reis, 1981; Olshavsky, 1976-77; Lundeberg, 1987), and familiarity with the topic (Pritchard, 1987). More research should be done to examine how reading processes differ across individuals and tasks. Populations of different age groups, readers of less proficiency, and expository texts all are reasonable extensions of this particular investigation. It would also be worthwhile to see how other reading tasks affect reading processes, such as reading directions to make

something, reading mysteries, or reading newspaper and magazine articles.

Further research is needed on the differences between retrospective and concurrent verbal reports. Although this research did not undertake such an examination, the researcher felt that a great deal was lost in the retrospective interviews, such as visualizing the text, paraphrasing, tieing the text in with prior knowledge, and expressing opinions of the text. What was reported in the retrospective interviews always agreed with what was reported in the concurrent interviews, thus each adding evidence of validity to the other. However, research studies that have only used retrospective verbal reporting may be losing a great deal of information.

Pedagogy

Educators need to be aware that different reading tasks and thus different reading tests may alter reading processes. In order to interpret test scores and to use these test scores wisely in making educational decisions, educators must be aware of what the tests are and are not measuring.

Having subjects describe what they are thinking and doing as they read a text and reporting

retrospectively what they were thinking and doing has several potentials in pedagogy for both students and teachers. The subjects in this study seemed to become more aware of all the available resources they had to comprehend the text as they used these techniques to express their introspective processes. While this may be a limitation to this study because subjects seemed to get better at relaying their processes as time went on, it may have helped them become more aware of their reading processes which could have a positive effect on their reading comprehension.

The researcher learned a great deal about reading processes and about each subject as a reader during these interviews. It is the opinion of the researcher that conducting similar interviews with students in the classroom would be helpful to any teacher in the diagnosis and assessment of reading. Teachers may also want to acquaint themselves with the literature on the values of modeling reading processes to students.

References

Adams, M.; Collins A. (1977). A schema-theoretic view of reading. Boston: Bolt, Beranek, and Newman.

Afflerbach, P.; Johnston, P. (1984). On the use of verbal reports in reading research. Journal of Reading Behavior, 16, 307-322.

Alvermann, D.; Ratekin, N. (1982). Metacognitive knowledge about reading proficiency: Its relation to study strategies and task demands. Paper presented at the Annual Meeting of the American Educational Research Association. New York, March 19-23, 1982.

Anderson, B. (1982). Test use today in elementary and secondary schools. In A.K. Wigdor and W.R. Garner (Eds.) Ability testing: Uses, consequences, and controversies, Part 2. Washington, DC: National Academy Press.

Ashby-Davis, C. (1985). Cloze and comprehension: A qualitative analysis and critique. Journal of Reading, 28, 585-89.

Bakan, D. (1954). A reconsideration of the problem of introspection. Psychological Bulletin, 51.

Baker, L.; Brown, A. (1984). Metacognitive skills in reading. In P.D. Pearson (Ed.), Handbook of research on reading. New York: Longman. 353-394.

Baker, L.; Brown, A. (1980). Metacognitive skills and reading. Technical Report No. 188. Urbana, IL. Center for the Study of Reading.

Bartlett, F. (1958). Thinking: An experimental and social study. London: Allen & Unwin.

Bateman, D.; Frandsen, K.; Dedmon, D. (1964). Dimensions of lecture comprehension: A factor analysis of listening test items. Journal of Communication, 14, 183-189.

Beyer, B. (1984a.). Improving Thinking Skills - Defining the problem. Phi Delta Kappan, March, 1984.

Beyer, B. (1984b.). Improving Thinking Skills - Practical Approaches. Phi Delta Kappan, April, 1984.

Beyer, B. (1984c.). Practical approaches to teaching thinking skills in every classroom. ASCD Curriculum Study Institute, Feb.

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals, Handbook I, cognitive domain. David McKay Company, Inc.

Boring, E.G. (1953). A history of introspection. Psychological Bulletin, 50.

Bower, G. (1978). Experiments on story comprehension and recall. Discourse Processes, 1, 211-231.

Brown, A.L. (1981). Metacognition: The development of selective attention strategies for learning from text. In M.L. Kamil (Ed.), Directions in reading: Research and instruction. Thirtieth yearbook of the National Reading Conference. Washington, DC: National Reading Conference. 21-43.

Campione, J.C., & Brown, A. (1985) Dynamic assessment: One approach and some initial data. (Tech. Rep. No. 361). Urbana: University of Illinois, Center for the Study of Reading.

Canney, J., & Winograd, P. (1979). Schemata for reading and reading comprehension performance. (Tech. Rep. No. 120). Champaign-Urbana; University of Illinois, Center for the Study of Reading.

Chaffee, J. (1985). Viewing reading and writing as thinking processes. Paper presented at the Annual Meeting of the American Educational Research Association (69th. Chicago, Illinois, March 31 - April 4, 1985).

Clark, C. (1982). Assessing free recall. The Reading Teacher, 35, 434-39.

Cohen, A. (1984). On taking language tests: What the students report. Language Learning, 1 (1), 70-81.

Cohen, A. (1986). Mentalistic measures in reading strategy research: some recent findings. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (20th Anaheim, CA, March 3-8, 1986).

Cohen, A. C. & Hosenfeld, C. (1981). Some uses of mentalistic data in second language research. Language Learning, 31.

Cohen, J. (1971). Thinking. Chicago: Rand McNally & Co., p.5.

The College Board (1986). Degrees of Reading Power Handbook.

Collins, A.; Brown, J.; Larkin, K. (1980). Inference in text understanding. In Theoretical issues in reading comprehension, edited by R. Spiro, B. Bruce, and W. Brewer. Hillsdale, N.J.: Erbaum.

Devine, T. (1986). Teaching reading comprehension: From theory to practice. Newton, MASS., Allyn and Bacon, Inc.

Dewey, J. (1910). How we think. Boston: D.C. Heath, 1910.

Dodge, R. (1912). The theory and limitations of introspection. American Journal of Psychology, 23.

Duffy, G.G.; Roehler, L.R.; Meloth, M.; Vavrus, L. (1988). Developing and evaluating measures associated with strategic reading. Journal of Reading Behavior.

Ericsson, K.; Simon, H. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Ericsson, A.; Simon, H. (1984). Protocol analysis. Campridge: Massachusetts Institute of Technology Press.

Farr, R. (1986) A response to Hoopfer and Hunsberger. FORUM in Reading and Language Education. Bloomington, IN: Indiana University, 129-133.

Farr, R.; Carey, R. (1986). Reading: What can be measured? Newark, Delaware: International Reading Association.

Farr, R. & Roser N. (1979). Teaching a child to read. New York: Harcourt Brace Jovanovich, Inc.

Flavell, J. (1976) Metacognitive aspects of problem solving. In The nature of intelligence. L. Resnick (Ed.). Hillsdale, NJ: Lawrence Erlbaum.

Flavell, J.; Wellman, H. (1977). Metamemory. In R. Kail & J. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Frannson, A. (1984). Cramming or understanding? Effects of intrinsic and extrinsic motivation on approach to learning and test performance. In J. C. Alderson & A. H. Urquhart (eds.), Reading in a Foreign Language. London: Longman, 86-121.

Garner, R. (1982). Verbal-report data on reading strategies. Journal of Reading Behavior, 14, 159-167.

Garner, R.; Reis, R. (1981). Monitoring and resolving comprehension obstacles: An investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. Reading Research Quarterly, 16, 569-582.

Goodman, K. S. (1967) Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 4, 126-35.

Goodman, Y; Burke, C. (1971). Reading miscue inventory--manual. New York: Macmillian.

Guilford, J. (1967). The nature of human intelligence. New York: McGraw-Hill.

Heap, J. (1982). What Counts as Reading: Limits to certainty in assessment. In R. Heyman (ed.)., The interactional context of the classroom: Talk and text. Calgary: The University of Calgary.

Henle, M. (1966). Cognitive Skills. In Learning about learning: A conference report, J. Bruner (Ed.). Washington, D.C.: U.S. Department of Health, Education and Welfare, 1966, pp. 53-63.

Hoffman, B. (1962). The tyranny of testing. New York: Collier.

Hoopfer, L.; Hunsberger, M. (1986). An ethnomethodo logical perspective on reading assessment. Forum in Reading and Language Education. Bloomington, IN., Indiana University Press.

Hopkins, K. D., George, C.A., & Williams, D. D. (1985). The concurrent validity of standardized achievement tests by content area using teachers' ratings as criteria. Journal of Educational Measurement, 22, 177-182.

Houts, P. (1977). The myth of measurability. New York: Hart.

Huey, E. (1908). The psychology and pedagogy of reading. New York: Macmillan. Reprinted by MIT Press, Cambridge, MA, 1968.

Hynd, C.; Alvermann, D. (1984). Verbal report data: The effect of manipulating text and familiarity and interval of reporting. Journal of Reading Education, 10, 1-7.

Irwin, P. & Mitchell, J. (1983). A procedure for assessing the richness of retellings. Journal of Reading, 26, 391-96.

Johnson, P. (1984). The expert mind: A new challenge for the information scientist. In T.M.A. Bemelmans (Ed.), Beyond productivity: Information systems development for organization effectiveness. Amsterdam: North Holland Publishing.

Johnson, R. E. (1970). Recall of prose as a function of the structural importance of the linguistic units. Journal of Verbal Learning and Verbal Behavior, 9.

Johnston, P. (1983). Reading comprehension assessment: A cognitive basis. Newark, DE: International Reading Association.

Johnston, P. (1986a). Assessing the process and the process of assessment in the language arts. In J. Squire (Ed.), The dynamics of language learning: Research in the language arts. Urbana, IL: National Council of Teachers of English.

Johnston, P. (1986b). Steps toward a more naturalistic approach to the assessment of the reading process. In J. Algina (Ed.), Advances in content-based educational assessment. New York: Ablex.

Kail, R. V., & Bisanz, J. (1982). Cognitive strategies. In C. R. Puff (Ed.), Handbook of research methods in human memory and cognition. New York: Academic Press.

Kalmbach, J. (1986). Getting at the point of retellings. Journal of Reading, 29, 326-33.

Kanfer, F. H. (1968). Verbal conditioning: A review of its current status. In T. R. Dixon & D. L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs, NJ: Prentice-Hall.

Kirkland, M. (1971). The effects of tests on students and schools. Review of Educational Research, 41 (4), 303-350.

Levine, M. (1976). The academic achievement test: Its historical context and social functions. American Psychologist, March, 1976.

Lundeberg, M. (1985). Studying understanding in legal case analysis (Doctoral dissertation, University of Minnesota, 1985). Dissertation Abstracts International, 46, 2630A.

Lundeberg, M. (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. Reading Research Quarterly, 22, 407-432.

MacKay, R. (1974). Standardized tests: objective/ob jectified measures of competence. In A. Cicourel, Language use and school performance. New York: Academic Press, Inc.

Markman, E. (1977). Realizing that you don't understand; A preliminary investigation. Child Development, 46, 986-992.

Markman, E. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child Development, 50, 643-655.

Mehan, H.; Wood, H. (1975). The Reality of ethnomethodoloy. New York: John Wiley and Sons.

Meichenbaum, D., Burland, S., Gruson, L., & Cameron, R. (1985). Metacognitive assessment. In S. R. Yussen (Ed.), The growth of reflection in children. Orlando: Academic Press.

National Assessment of Educational Progress (1981). Reading, thinking, and writing: Results from the 1979-80 national assessment of reading and literature. Education Commission of the States.

Nickerson, R. (1981) Thoughts on teaching thinking. Educational Leadership, 39(1).

Nisbett, R; Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Olshavsky, J. (1976-77). Reading as problem-solving: An investigation of strategies. Reading Research Quarterly, 12, 654-674.

Paris, S. G.; Lindauer, B. K. (1976). The development of cognitive skills during childhood. In B.W. Wolman (Ed.) Handbook of developmental psychology. Englewood Cliffs, NJ: Prentice-Hall.

Paris, S.G.; Lipson, M; Wixson, K. (1983). Becoming a strategic reader. Contemporary Educational Psychology, 8, 293-316.

Pearson, D.; Spiro, R. (1980). Toward a theory of reading comprehension instruction. Topics in Language Disorders, 1, 71-88.

Pearson, D. (1986). Assessment, accountability, and professional prerogative. Presidential address at the National Reading Conference, Austin, Texas.

Pepper, S. C. (1918). What is introspection? American Journal of Psychology, 29.

Polanyi, M. (1973). The tacit dimension. Garden City, NY: Doubleday.

Presseisen, B. (1984) Thinking skills: Meanings, models, and materials. Philadelphia, PA.: Research for Better Schools, Inc.

Pressley, M.; Forrest-Pressley, D.; Elliott-Faust, D.; Miller, G. (1985). Children's use of cognitive strategies, how to teach strategies and what to do if they can't be taught. In M. Pressley & C. Brainerd (Eds.), Cognitive Learning and Memory in Children: New York: Springer-Verlag, 1-47.

Pritchard, R. (1987). A cross-cultural study of the effects of cultural schemata on proficient readers' comprehension monitoring strategies and their comprehension of culturally familiar and unfamiliar passages. Unpublished Doctoral Dissertation, Indiana University.

Rosenblatt, L. (1978). The Reader, the text, the poem. University Press Carbondale.

Ross, A. (1983). Alan Ross' comments. In D. Cornine, D. Elkind, A. D. Hendrickson, D. Meichenbaum, R. L. Sieber & F. Smith (Eds.), Interdisciplinary voices in learning disabilities and remedial education. Austin, TX: Pro ed.

Roth, D. (1974). Intelligence testing as a social activity. In A. Cicourel, Language use and school performance. New York: Academic Press, Inc.

Rowe, D. W., and Rayford, L. (1984). Instantiating background knowledge in reading comprehension assessment. Paper presented at the 34th annual National Reading Conference, St. Petersburg, FL.

Royer, J. & Cunningham, D. (1978). On the theory and measurement of reading comprehension. (Tech. Rep. No. 91). Urbana: University of Illinois, Center for the Study of Reading, June, 1978. (ED 157 040).

Rumelhart, D. (1975). Notes on a Schema for Stories. In D. Bobrow & A. Collins (Eds.), Representation and understanding studies in cognitive science: New York: Academic Press.

Singer, H. (1983). The substrata-factor theory of reading: Its history and Conceptual relationship to interaction theory. In L. Gentile; M. Kamil, & J. Blanchard (Eds.), Reading research revisited: Columbus, Ohio: Merrill.

Smith, E. R., & Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. Psychological Review, 85.

Smith, S. (1980). Assessing student learning and retention strategies using retellings of text passages. Report of the Learning Skills Center. Bloomington, Ind.: Indiana University.

Smith, S., & Jackson, J. (1985). Assessing reading/learning skills with written retellings. Journal of Reading, 28, 622-31.

Sprio, R. J. (1980). Schema theory and reading comprehension: New directions (Technical Report No. 191). Urbana: University of Illinois, Center for the Study of Reading.

Spiro, R.J. & Meyers, A. (1984). Individual differences and underlying cognitive processes. In Handbook of reading research (pp. 471-504). New York: Longman.

Sternberg, R. J. (1985). Review of Meichenbaum, Burland, and Cameron's metacognitive assessment. In S. R. Yussen (Ed.), The growth of reflection in children. Orlando: Academic Press.

Strenio, A., Jr. (1981). The testing trap. New York: Rawson, Wade Publishers, Inc.

Taylor, W. L. (1953) Cloze procedure: A new tool for measuring readability. Journalism Quarterly 30, Fall, 415-33.

Thorndike, E. (1917). Reading as reasoning: A study of mistakes in paragraph reading. Journal of Reading Behavior, 8, 323-332.

Titchener, E. B. (1912a). Protegomena to a study of introspection. American Journal of Psychology, 23.

Titchener, E. B. (1912b). The schema of introspection. American Journal of Psychology, 23.

Valencia, S.; Pearson, D. (1986). New models for reading assessment. Center for the Study of reading.

Valencia, S.; Pearson, D. (1987). Reading assessment: time for a change. The Reading Teacher, 40 (8), 726-732.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

Watson, J. B. (1913). Psychology as the behaviorist views it. Psychological Review, 20.

White, P. (1980). Limitations of verbal reports of internal events: A refutation of Nisbett and Wilson and of Bem. Psychological Review, 87, 105-112.

Wingenbach, N. (1984). The gifted reader: Metacognition and comprehension strategies. Paper presented at the Annual Meeting of the National Council of Teachers of English Spring Conference (3rd), Columbus, OH, April 12-14, 1984.

Winograd, T. (1972). Understanding natural language. Cognitive Psychology, 3, 1-191.

Wixson, K; Peters, C. (1984). Reading redefined: a Michigan Reading Association position paper. The Michigan Reading Journal, 17, 4-7.

APPENDIX

FRIEDMAN TWO-WAY ANOVA

	MEAN RANK	VARIABLE
RECALL	3.50	Multiple Choice
	2.00	Cloze Test
	3.28	Written Retelling
	1.22	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
18.8333	3	.0003

	MEAN RANK	VARIABLE
KEY WORD	3.61	Multiple Choice
	1.78	Cloze Test
	2.83	Written Retelling
	1.78	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
12.9000	3	.0049

	MEAN RANK	VARIABLE
ELIMINATE	3.78	Multiple Choice
	2.56	Cloze Test
	1.83	Written Retelling
	1.83	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
13.6333	3	.0034

	MEAN RANK	VARIABLE
GUESSING	2.44	Multiple Choice
	4.00	Cloze Test
	1.89	Written Retelling
	1.67	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
17.933	3	.0005

	MEAN RANK	VARIABLE
MAIN IDEA	2.50	Multiple Choice
	1.89	Cloze Test
	3.61	Written Retelling
	2.00	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
10.0333	3	.0183

	MEAN RANK	VARIABLE
SKIP	2.67	Multiple Choice
	3.78	Cloze Test
	1.78	Written Retelling
	1.78	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
14.6000	3	.0022

	MEAN RANK	VARIABLE
CONTEXT CLUES	2.00	Multiple Choice
	4.00	Cloze Test
	2.00	Written Retelling
	2.00	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
16.2000	3	.0010

	MEAN RANK	VARIABLE
REREAD	2.94	Multiple Choice
	3.72	Cloze Test
	2.28	Written Retelling
	1.06	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
20.6667	3	.0001

	MEAN RANK	VARIABLE
PARAPHRASE	2.61	Multiple Choice
	1.50	Cloze Test
	2.50	Written Retelling
	3.39	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
9.7333	3	.0210

	MEAN RANK	VARIABLE
OPINION	2.67	Multiple Choice
	1.61	Cloze Test
	1.94	Written Retelling
	3.78	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
14.9000	3	.0019

	MEAN RANK	VARIABLE
VISUALIZE	2.83	Multiple Choice
	1.44	Cloze Test
	2.28	Written Retelling
	3.44	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
11.7000	3	.0085

*	MEAN RANK	VARIABLE
READ AHEAD	2.17	Multiple Choice
	3.50	Cloze Test
	2.17	Written Retelling
	2.17	Nonassessed Task

CHI-SQUARE	D.F.	SIGNIFICANCE
7.2000	3	.0658

*Friedman Test failed to show significance due to ranking, but post-hoc analysis confirmed statistical significance.

JANET L. POWELL

EDUCATION

Doctor of Education, Language Education, Elementary Reading Emphasis with a cognate in Elementary School Administration, Indiana University, December, 1988

Master of Science, Language Education, Elementary Reading Emphasis, Indiana University, May, 1984

Bachelor of Science, Elementary Education, Reading Endorsement, Indiana University, December, 1977

PROFESSIONAL EXPERIENCE

Assistant Professor, Ball State University, Muncie, IN, 1988-present

Associate Instructor, Indiana University, Bloomington, IN, 1983-1988

Consultant, Harcourt Brace Jovanovich, Orlando, FL, 1986-1988

Elementary Education Teacher, Monroe County Community Schools, Bloomington, IN, 1984-85; 1978-82